REVIEW Open access

Ensemble Methods in Materials AI — Why Models Disagree and What It Means Scientifically

Claire Martin[1*], Julien Robert[2], Sophie Bernard[1]

Abstract

The integration of artificial intelligence (AI) into materials science has revolutionized how we predict, design, and discover new materials. Among various AI techniques, ensemble methods have emerged as powerful tools that leverage the collective intelligence of multiple models to enhance prediction accuracy and reliability. This review explores the application of ensemble methods in materials AI, focusing on why individual models disagree and the scientific implications of such disagreements. By analyzing recent advancements, we highlight how ensemble approaches address uncertainties in material property prediction, phase stability, and electronic structure calculations. The review synthesizes insights from peer-reviewed literature published, emphasizing the role of ensemble methods in providing robust predictions and uncovering underlying physical principles. Ultimately, understanding model disagreement not only improves computational efficiency but also deepens our scientific understanding of material behavior.

Keywords Materials AI, Property prediction, Uncertainty quantification, Ensemble methods, Model disagreement, Materials design

*Correspondence:
Claire Martin
claire.martin@gmail.com
[1] Department of Computational Materials Research, Faculty of Engineering, University of Lyon, Lyon, France
[2] Department of Artificial Intelligence Systems, Faculty of Engineering, University of Strasbourg, Strasbourg, France

Introduction

Materials science has historically advanced through a close interplay between experimental investigation and theoretical modeling, with progress often driven by incremental refinement of synthesis techniques and physical theories. While this paradigm has yielded transformative materials—from semiconductors to structural alloys—it faces increasing limitations in the modern era. Contemporary material systems are frequently characterized by high compositional complexity, multiscale interactions, and strongly nonlinear structure–property relationships. Combined with the immense size of chemical and structural design spaces, these factors render traditional trial-and-error experimentation and first-principles calculations increasingly time-consuming, computationally expensive, and resource-intensive [1-3]. As a result, there is a growing need for approaches that can accelerate materials discovery while maintaining predictive reliability.

The emergence of machine learning (ML) has introduced a paradigm shift in how materials are designed, analyzed, and optimized. By leveraging large datasets generated from experiments, simulations, and high-throughput workflows, ML models can uncover hidden patterns and correlations that are difficult to capture using conventional approaches [1, 3, 4-9]. In materials informatics, ML has been successfully applied to a wide range of tasks, including property prediction, phase stability assessment, defect characterization, and inverse materials design. These data-driven methods enable

rapid screening of candidate materials and significantly reduce the cost and time required to identify promising compositions or structures.

Within the broader landscape of ML techniques, ensemble methods occupy a particularly important position. Rather than relying on a single predictive model, ensemble approaches combine multiple base learners to produce aggregated predictions that are often more accurate and robust than those of individual models [2, 4]. Common ensemble strategies—such as bagging, boosting, and stacking—exploit model diversity to reduce variance, bias, or both, depending on the specific formulation [2, 4]. In the context of materials AI, ensemble methods have demonstrated strong performance across diverse applications, including phase prediction in high-entropy alloys, mechanical property estimation of cementitious composites, and microstructure–property mapping in complex materials systems [2, 10-16].

The appeal of ensemble methods in materials science stems largely from their ability to cope with uncertainty and data limitations. Experimental datasets in materials research are often noisy, incomplete, or heterogeneous due to variations in synthesis routes, measurement techniques, and reporting standards [5, 6, 17, 18]. Similarly, simulation-derived data may contain systematic errors arising from model assumptions or numerical approximations. By integrating multiple models with different inductive biases, ensemble methods can mitigate the impact of such uncertainties and yield more stable predictions, making them particularly well-suited for real-world materials problems.

A defining feature of ensemble learning is disagreement among base models. Such disagreement can arise from differences in training subsets, feature representations, model architectures, or hyperparameter choices [4, 7, 10, 18]. Importantly, this diversity is not merely a technical byproduct of model construction but a source of scientifically meaningful information. Regions where ensemble members disagree strongly often correspond to sparse data coverage, extrapolative regimes, or competing physical interpretations. Consequently, model disagreement can serve as a proxy for epistemic uncertainty, highlighting areas where predictions should be treated with caution or where additional data collection is warranted [4, 10, 12, 17, 19-22].

Beyond uncertainty estimation, ensemble disagreement can also contribute to knowledge discovery. In electronic structure modeling, for example, discrepancies in predictions across different density functional theory (DFT) approximations reflect underlying theoretical uncertainties. Ensemble ML approaches can reconcile these differences by learning consensus trends, thereby improving predictive accuracy while also exposing systematic biases in the underlying methods [10]. More broadly, analyzing disagreement patterns may reveal limitations of existing descriptors, suggest missing physical variables, or even point toward previously unrecognized phenomena in complex materials systems [4, 12, 17].

Against this backdrop, the objectives of this review are threefold. First, it aims to provide a comprehensive overview of ensemble learning methods and their current applications in materials-focused artificial intelligence. Second, it seeks to analyze the origins of model disagreement within ensembles, emphasizing both algorithmic and data-related factors. Third, it discusses the scientific implications of such disagreements, particularly their role in uncertainty quantification and their potential to facilitate insight-driven materials discovery. By synthesizing recent literature, this narrative review aims to guide researchers in the effective use of ensemble methods as tools not only for improved prediction but also for more reliable, interpretable, and insightful materials research.

Main Text

Overview of ensemble methods in materials AI

Ensemble methods are rooted in the central idea that a collection of relatively simple or moderately accurate models—often referred to as weak learners—can be combined to form a predictive system that outperforms any individual constituent model [2, 4, 23-25]. This improvement arises from exploiting diversity among learners, which enables ensembles to counteract overfitting, reduce prediction variance, and, in some cases, correct systematic bias. In the field

of artificial intelligence for materials, where datasets are frequently heterogeneous and incomplete, this principle has proven particularly valuable.

Among the most widely adopted ensemble techniques in materials AI are random forests, gradient boosting machines, and adaptive boosting frameworks. These approaches have been successfully applied to predict a broad range of material properties, including compressive strength, hardness, elastic moduli, and phase stability [2, 16, 24, 26–30]. Random forests, which are based on the bagging paradigm, construct an ensemble of decision trees trained on bootstrapped subsets of the data while introducing randomness in feature selection. This strategy decorrelates individual trees and yields robust predictions, even in the presence of noisy or nonlinear data. For example, random forest models have been employed to predict material hardness by aggregating the decisions of multiple trees, each capturing different aspects of the structure–property relationship [25].

Boosting-based ensembles, such as gradient boosting and adaptive boosting, follow a fundamentally different philosophy. Rather than training models independently, boosting methods sequentially construct learners that focus on correcting the errors of previous models. This makes them particularly effective for capturing subtle trends in complex datasets. In materials science applications, boosting ensembles have demonstrated strong performance in regression tasks involving highly nonlinear relationships, such as those encountered in composite materials and multi-phase systems [16, 27, 28]. However, their sensitivity to noisy labels also underscores the importance of careful data preprocessing in materials datasets.

In parallel with classical ML ensembles, deep learning ensembles have gained increasing attention in materials science [1, 4, 18, 28]. These approaches typically involve training multiple neural networks with different random initializations, architectures, or training subsets, and then averaging or otherwise aggregating their outputs [4, 18]. While individual neural networks can be prone to overconfidence, especially in extrapolative regimes, ensemble averaging has been shown to improve both predictive accuracy and uncertainty calibration significantly. In atom-resolved microscopy, for instance, ensemble neural networks combined with iterative training schemes have enabled uncertainty-aware image interpretation and guided automated experimental workflows [4]. Similarly, ensembles of message passing neural networks have been applied to molecular and materials datasets to provide calibrated uncertainty estimates alongside property predictions [28]. The ensemble strategies employed in materials AI can be systematically categorized according to their learning philosophy and scientific role **(Table 1)**.

Table 1. Ensemble learning paradigms in materials artificial intelligence and their methodological roles

Ensemble paradigm	Core mechanism	Typical base learners	Materials AI use cases	Strengths	Limitations
Bagging (e.g., random forests)	Parallel training on bootstrapped datasets	Decision trees	Hardness prediction, phase classification, microstructure–property mapping	Variance reduction, robustness to noise	Limited bias correction
Boosting (e.g., gradient boosting, adaboost)	Sequential error correction	Trees, weak regressors	Mechanical property regression, composite materials	Captures nonlinear trends, high accuracy	Sensitive to noisy labels
Stacking	Meta-learner combines heterogeneous models	SVMs, NNs, trees	Phase stability, multi-property prediction	Exploits model complementarity	Increased complexity

Deep ensembles	Multiple neural networks with varied initialization/architecture	CNNs, GNNs, MPNNs	Atomistic modeling, microscopy, and molecular properties	Calibrated uncertainty, expressive	Computational cost
Theory-aware ensembles	Models trained on diverse theoretical outputs	ML surrogates of DFT	Electronic structure prediction	Reconciles theoretical disagreement	Depends on the quality of the theory

The growing adoption of ensemble methods in materials AI is closely tied to the field's unique data challenges. Materials datasets are often small relative to the dimensionality of the feature space, and collecting new experimental data can be expensive or impractical [5, 6]. To address these constraints, small-data ML strategies—such as transfer learning, data augmentation, and ensemble learning—have been proposed as complementary solutions [5]. At the same time, in domains where large datasets are available, such as porous materials and materials genomics, ensemble methods have been integrated into big-data pipelines to improve robustness and generalization in high-throughput materials design [21]. Together, these developments highlight the flexibility of ensemble learning across both data-scarce and data-rich regimes.

Applications in property prediction and materials design

One of the most mature application areas for ensemble methods in materials AI is property prediction, where ensembles consistently outperform single-model approaches in terms of accuracy and reliability [2, 16, 24, 25, 27–30]. In the context of high-entropy alloys, ensemble-based frameworks have been used to predict phase formation by integrating outputs from diverse ML models, including support vector machines, neural networks, and decision trees [2]. This multi-model integration is particularly advantageous for high-entropy systems, where complex compositional interactions challenge traditional phase prediction methods.

Similarly, ensemble learning has played a key role in the discovery and screening of superhard materials. By combining predictions from multiple algorithms, ensemble approaches reduce the risk of false positives and improve confidence in candidate selection, thereby accelerating materials discovery pipelines [25]. These capabilities are especially valuable when experimental validation is costly or time-intensive.

In cementitious and construction-related materials, ensemble methods have been extensively applied to predict compressive strength and other mechanical properties using compositional and processing variables [16, 24, 27–30]. Such applications are closely aligned with sustainability goals, as they enable the incorporation of industrial waste materials into concrete formulations while maintaining performance standards. Gradient boosting and bagging-based ensembles have repeatedly demonstrated superior predictive performance compared to individual models in estimating the strength of high-performance and eco-friendly concrete systems [28]. Beyond regression tasks, ensemble classifiers have also been used to identify failure modes and predict bearing capacity in coal-grout materials and reinforced concrete columns, supporting safer and more efficient structural design [29, 30].

Beyond traditional structural materials, ensemble methods have increasingly contributed to the design of functional materials, including thermoelectric, electrocatalytic, and biomedical materials [12, 19, 24]. In electrocatalysis, ensemble-based ML frameworks have accelerated the discovery of low-dimensional catalysts for hydrogen evolution reactions by efficiently navigating large chemical design spaces [24]. In glass-based materials for healthcare applications, ensemble models have been employed to control dissolution behavior, enabling the rational design of bioactive compositions with tailored degradation rates [12]. These examples illustrate how ensemble learning extends beyond prediction accuracy to influence practical materials design and optimization.

Collectively, these applications demonstrate that ensemble methods serve not only as performance-enhancing tools but also as enablers of reliability and interpretability in materials AI. By integrating diverse modeling perspectives, ensembles help bridge the gap between data-driven predictions and the complex physical realities of materials systems, laying the groundwork for more trustworthy and impactful AI-assisted materials research.

Reasons for model disagreement in ensembles

Model disagreement within ensemble learning frameworks arises from a combination of data-related, algorithmic, and fundamentally scientific factors [4, 7, 10, 18]. In materials AI, these sources of disagreement are especially pronounced due to the intrinsic complexity of materials systems and the limitations of available datasets. Rather than representing failure, such disagreements reflect the interaction between model assumptions, data coverage, and underlying physical phenomena.

Data quantity and quality constitute a primary source of disagreement among ensemble members. Sparse datasets, measurement noise, and inconsistencies across experimental protocols can cause models trained on different subsets of data to converge toward distinct predictive hypotheses [6]. This issue is pervasive in materials science, where experimental campaigns are costly, and datasets are often biased toward specific material classes or processing conditions. Ensemble methods expose these limitations by exhibiting increased prediction variance in poorly sampled regions of the feature space, thereby offering a quantitative indicator of data-driven uncertainty [5, 18].

A second, deliberately engineered source of disagreement is model diversity. Ensemble learning explicitly encourages heterogeneity among base learners by varying training data, model architectures, feature representations, or optimization strategies [2, 4]. This diversity enables ensembles to capture complementary perspectives on complex structure–property relationships. For example, in uncertainty quantification for material property prediction, different approaches—such as Monte Carlo dropout and deep neural network ensembles—can produce divergent uncertainty estimates, each reflecting distinct assumptions about model confidence and data noise [18]. In electronic structure modeling, ML-based roadmaps have emphasized that ensembles trained on outputs from different density functional theory (DFT) approximations can learn from inter-method discrepancies, effectively reconciling divergent theoretical viewpoints [10]. The primary drivers of ensemble disagreement in materials AI span data, algorithmic, and physical dimensions, each carrying distinct epistemic interpretations **(Table 2)**.

Table 2. Origins of model disagreement in ensemble materials AI and their epistemic interpretation

Source of disagreement	Origin	Manifestation in ensembles	Scientific meaning	Typical signals
Data sparsity	Limited experimental coverage	High variance across models	Epistemic uncertainty	Wide prediction spread
Measurement noise	Experimental inconsistency	Unstable predictions	Aleatoric uncertainty	Inconsistent confidence
Model inductive bias	Algorithmic assumptions	Divergent trends	Competing hypotheses	Systematic disagreement
Feature representation	Descriptor incompleteness	Sensitivity to inputs	Missing physics	Boundary instability
Physical regime transitions	Phase changes, correlations	Localized disagreement	Regime shift	Phase-boundary divergence

Beyond algorithmic considerations, intrinsic scientific characteristics of materials systems can also drive ensemble disagreement. Phenomena such as strong electronic correlation, multireference character in quantum chemical calculations, and composition-driven phase transitions introduce abrupt changes in property landscapes that are difficult for data-driven models to capture uniformly [2, 10]. In such cases, ensembles may exhibit systematic disagreement near phase boundaries or in regions dominated by competing physical mechanisms. These patterns of disagreement have been leveraged to define domains of applicability, distinguishing regions where model predictions are reliable from those

where extrapolation or unmodeled physics dominate [7]. Consequently, disagreement can serve as an early warning signal of unexplored physical regimes or the breakdown of learned correlations [7, 17].

Scientific implications of model disagreement

Crucially, model disagreement in ensembles should not be viewed as a limitation, but rather as a powerful source of scientific information [4, 10, 18]. One of its most immediate and practical benefits lies in robust uncertainty quantification, which is essential for trustworthy deployment of ML models in materials research. Ensemble-based uncertainty estimates have been shown to improve the calibration of predictions, enabling researchers to distinguish between high-confidence and low-confidence predictions in molecular and materials systems [4, 18, 28]. For example, neural network ensembles have yielded more reliable uncertainty bounds for property prediction tasks, supporting risk-aware decision-making in materials screening workflows [28]. Rather than being suppressed, ensemble disagreement can be operationalized to support uncertainty estimation, experimental steering, and theory refinement **(Table 3)**.

Table 3. How ensemble disagreement contributes to scientific inference in materials AI

Scientific function	How disagreement is used	Benefit	Example domains
Uncertainty quantification	Variance across ensemble predictions	Risk-aware screening	Property prediction
Domain of applicability	Identifying low-consensus regions	Prevents over-extrapolation	Materials discovery
Experimental prioritization	Sampling high-disagreement regions	Efficient data acquisition	Autonomous labs
Theory reconciliation	Learning consensus from conflicting theories	Robust predictions	DFT-based discovery
Hypothesis generation	Analyzing divergence patterns	New physical insight	Structure analysis

Beyond uncertainty estimation, ensemble disagreement can actively guide experimental and computational workflows. In atom-resolved microscopy, disagreement among ensemble predictions has been used to prioritize regions of interest for further data acquisition, enabling iterative learning cycles that improve both model performance and experimental efficiency [4]. Such closed-loop frameworks illustrate how disagreement can be transformed into actionable insights that directly influence experimental design.

At a deeper scientific level, patterns of disagreement can reveal limitations in existing theoretical models and illuminate underlying physical principles. In DFT-driven materials discovery, ensemble ML approaches have been used to mitigate inconsistencies across different exchange–correlation functionals, yielding consensus predictions that are more robust than any single approximation [10]. This strategy not only improves predictive performance but also provides insight into the sensitivity of material properties to electronic structure approximations, with implications for understanding stability, bonding, and electronic behavior in complex materials [10, 17].

Furthermore, model disagreement catalyzes knowledge discovery. By systematically analyzing where and why ensemble members diverge, researchers can identify deficiencies in training data, missing descriptors, or inadequacies in current physical models [8, 13, 20]. In X-ray diffraction analysis, for instance, interpretable ensemble methods have uncovered latent symmetry relationships, enhancing the characterization of crystalline materials and revealing structural motifs that were not readily apparent using conventional analysis techniques [20]. More broadly, embracing disagreement within ensembles encourages a shift from purely predictive modeling toward hypothesis generation, enabling ML to function as a tool for scientific exploration rather than a black-box predictor [3, 9, 15, 23].

In this sense, ensemble disagreement occupies a dual role in materials AI: it enhances reliability through uncertainty-aware prediction while simultaneously opening pathways to deeper scientific understanding. When properly interpreted, disagreement becomes a signal—pointing not to failure, but to opportunity—thereby accelerating innovation and advancing the fundamental understanding of materials behavior.

Results and Discussion

The systematic exploration of ensemble methods in materials artificial intelligence reveals a nuanced landscape in which model disagreement functions simultaneously as a methodological challenge and a catalyst for scientific progress [1–30]. A central insight emerging from this body of work is that ensembles do not simply suppress errors through averaging; rather, they actively surface epistemic uncertainties that are intrinsic to materials modeling, data availability, and theoretical assumptions [4, 10, 18]. This distinction is particularly important in materials science, where incomplete knowledge and sparse observations are the norm rather than the exception. The epistemic role of ensemble disagreement as a bridge between modeling uncertainty and scientific inference is summarized schematically in **Figure 1**.

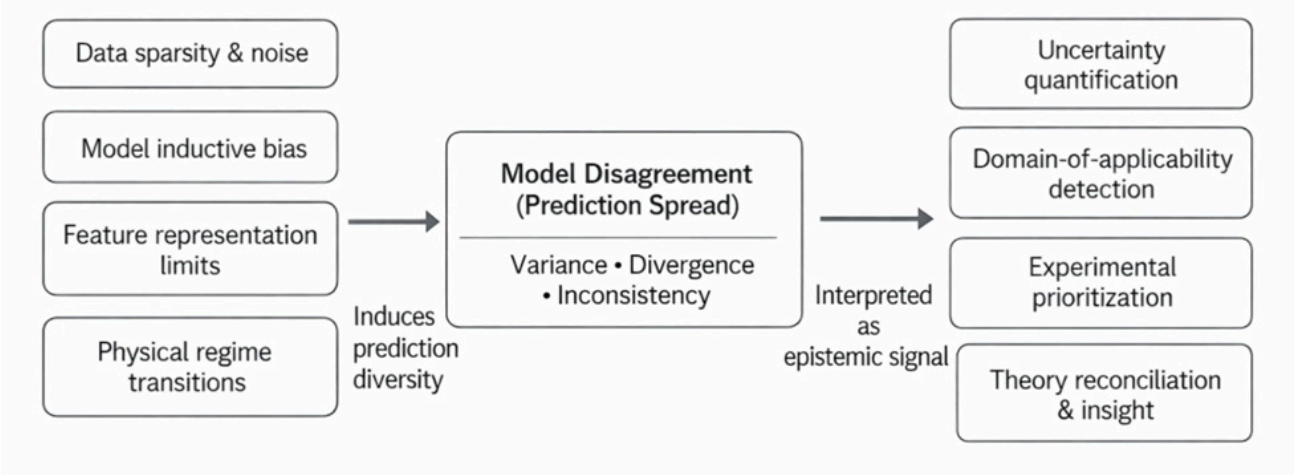

Figure 1. Conceptual role of model disagreement in ensemble materials artificial intelligence

In data-limited regimes—common across experimental materials research—ensemble methods mitigate overfitting by promoting diversity in learned representations [5, 6]. However, the resulting disagreements among ensemble members often serve as diagnostic signals, revealing deficiencies in data coverage, biases in sampling, or sensitivity to noise that would remain hidden in single-model approaches [5, 7, 13]. In property-prediction tasks, such as mechanical-strength or phase-stability estimation, ensemble variance has proven effective at capturing aleatoric uncertainty arising from experimental variability and measurement noise [18, 28]. This capacity to expose uncertainty is especially valuable for decision-making in materials design, where false confidence can lead to costly experimental failures.

Beyond uncertainty quantification, the scientific significance of ensemble disagreement extends to theory refinement and physical insight. In electronic structure applications, ensemble learning has been used to reconcile discrepancies among different density functional theory (DFT) approximations, suggesting that persistent disagreements are symptomatic of incomplete treatments of electron correlation and exchange effects [10]. Rather than selecting a single “best” functional, ensemble-based consensus approaches acknowledge theoretical plurality and exploit it to improve robustness. Similarly, in symmetry and structure prediction from X-ray diffraction data, interpretable ensemble models have revealed latent structural patterns, demonstrating how disagreement can expose hidden regularities in complex datasets [20]. These examples underscore the role of ensembles as tools not only for prediction, but also for hypothesis generation and theory interrogation.

The ability of ensembles to identify domains of applicability represents another important scientific contribution. Disagreement patterns often delineate boundaries between interpolation and extrapolation regimes, providing early warnings against unreliable predictions [7, 23]. This capability is particularly relevant in materials discovery, where ML models are frequently deployed beyond the strict confines of their training data. By highlighting regions of low consensus, ensembles help prevent overconfident extrapolation and support more cautious, physically informed exploration of new material spaces.

Despite these strengths, significant challenges remain in interpreting and operationalizing ensemble disagreement. Techniques such as k-fold forward cross-validation can enhance exploratory prediction power in small datasets, but may also amplify disagreement in high-dimensional feature spaces, complicating interpretation [13]. Quantifying and comparing diversity across ensemble members remains a methodological challenge, especially when heterogeneous models or representations are involved [2, 4]. Moreover, while ensemble methods have demonstrated strong performance in interdisciplinary domains such as environmental engineering, renewable energy systems, and infrastructure materials [15, 26, 27], the lack of standardized uncertainty quantification frameworks limits reproducibility and cross-study comparison [18].

Addressing these challenges will likely require deeper integration between ensemble learning and physics-informed machine learning. Embedding physical constraints into ensemble members offers a promising route for interpreting disagreement: divergence among models may then signal either violations of physical assumptions or the presence of unmodeled phenomena [17, 22]. At the same time, ethical and practical considerations must be acknowledged. Ensemble methods often incur higher computational costs than single-model approaches, raising questions about scalability and efficiency [1, 9]. While ensembles scale effectively in data-rich contexts such as porous materials genomics [21], their application to atomistic simulations and quantum-level modeling demands careful algorithmic optimization [17, 22].

Looking ahead, hybrid strategies that combine ensembles with interpretable machine learning, symbolic regression, or causal inference frameworks offer a promising pathway toward demystifying model disagreement [8, 20]. Such approaches could transform disagreement from a statistical artifact into a transparent explanatory signal. Ultimately, embracing ensemble disagreement encourages a shift in perspective—from viewing AI as a black-box predictor to recognizing it as a scientific collaborator that challenges assumptions, highlights uncertainty, and stimulates discovery [3, 14, 19].

Conclusion

In conclusion, ensemble methods have established themselves as foundational tools in materials artificial intelligence, enhancing predictive robustness while offering deeper insight into the origins and implications of model disagreement. This review has shown that disagreement within ensembles arises from a confluence of data limitations, intentional model diversity, and intrinsic material complexity, each reflecting different dimensions of uncertainty and physical reality. Far from being a weakness, such disagreement provides a valuable lens through which uncertainty can be quantified, theoretical assumptions can be interrogated, and new scientific knowledge can be generated.

Scientifically, ensemble disagreement has enabled advances ranging from uncertainty-calibrated materials design to the discovery of novel structure–property relationships and candidate materials. By highlighting where models disagree, ensembles guide targeted data acquisition, refine modeling strategies, and prevent overconfident extrapolation. These attributes are increasingly critical as materials AI moves toward autonomous experimentation and closed-loop discovery systems.

Looking forward, continued progress will depend on improving the interpretability of ensemble methods and integrating them more tightly with experimental and simulation workflows. As datasets grow in size and diversity and ML methodologies continue to evolve, ensemble approaches are poised to play a central role not only in accelerating materials discovery, but also in deepening scientific understanding of why models diverge and what such divergence

reveals about the materials themselves. Researchers are therefore encouraged to leverage ensemble methods not merely for enhanced accuracy, but for the richer scientific narratives and insights that emerge when disagreement is embraced rather than suppressed.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 08 Feb 2023 Revised: 01 Apr 2023 Accepted: 10 May 2023
Published online: 18 July 2023

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Choudhary K, DeCost B, Chen C, Jain A. Recent advances and applications of deep learning methods in materials science. npj Comput Mater. 2022;8(1):59.
https://doi.org/10.1038/s41524-022-00734-6.

Mishra A, Kompella L, Sanagavarapu LM. Ensemble-based machine learning models for phase prediction in high entropy alloys. Comput Mater Sci. 2022;204:111074.
https://doi.org/10.1016/j.commatsci.2021.111074.

Morgan D, Jacobs R. Opportunities and challenges for machine learning in materials science. Annu Rev Mater Res. 2020;50:71-103.
https://doi.org/10.1146/annurev-matsci-070218-010015.

Ghosh A, Sumpter BG, Dyck O, Kalinin SV. Ensemble learning-iterative training machine learning for uncertainty quantification and automated experiment in atom-resolved microscopy. npj Comput Mater. 2021;7(1):140. https://doi.org/10.1038/s41524-021-00569-7.

Xu P, Ji X, Li M, Lu W. Small data machine learning in materials science. npj Comput Mater. 2023;9(1):42. https://doi.org/10.1038/s41524-023-01000-z.

Liu Y, Yang Z, Zou X, Ma S, Liu D. Data quantity governance for machine learning in materials science. Natl Sci Rev. 2023;10(7):nwad125.

Sutton C, Boley M, Ghiringhelli LM, Rupp M. Identifying domains of applicability of machine learning models for materials science. Nat Commun. 2020;11(1):6130. https://doi.org/10.1038/s41467-020-17112-9.

Rodrigues JF Jr, Florea L, De Oliveira MCF. Big data and machine learning for materials science. Discov Mater. 2021;1(1):12. https://doi.org/10.1007/s43939-021-00012-0.

Gao C, Min X, Fang M, Tao T, Zheng X. Innovative materials science via machine learning. Adv Funct Mater. 2022;32(18):2108044. https://doi.org/10.1002/adfm.202108044.

Kulik HJ, Hammerschmidt T, Schmidt J, Botti S. Roadmap on machine learning in electronic structure. Electron Struct. 2022;4(2):023004.

Batra R, Chan H, Kamath G, Ramprasad R. Screening of therapeutic agents for covid-19 using machine learning and ensemble docking studies. J Phys Chem Lett. 2021;12(32):7727-33. https://doi.org/10.1021/acs.jpclett.0c02278.

Han T, Stone-Weiss N, Huang J, Goel A, Kumar A. Machine learning as a tool to design glasses with controlled dissolution for healthcare applications. Acta Biomater. 2020;107:286-98. https://doi.org/10.1016/j.actbio.2020.02.027.

Xiong Z, Cui Y, Liu Z, Zhao Y, Hu M, Hu J. Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation. New J Chem. 2020;44(37):15865-75. https://doi.org/10.1039/D0NJ02985A.

Chen C, Zuo Y, Ye W, Li X, Deng Z. A critical review of machine learning of energy materials. Adv Energy Mater. 2020;10(8):1903242. https://doi.org/10.1002/aenm.201903242.

Zhong S, Zhang K, Bagheri M, Burken JG. Machine learning: New ideas and tools in environmental science and engineering. Environ Sci Technol. 2021;55(19):12741-54. https://doi.org/10.1021/acs.est.1c01339.

Marani A, Nehdi ML. Machine learning prediction of compressive strength for phase change materials integrated cementitious composites. Constr Build Mater. 2020;265:120286. https://doi.org/10.1016/j.conbuildmat.2020.121286.

Axelrod S, Schwalbe-Koda D. Learning matter: Materials design with machine learning and atomistic simulations. Acc Mater Res. 2022;3(3):343-57. https://doi.org/10.1021/accountsmr.1c00238.

Tran K, Neiswanger W, Yoon J, Zhang Q. Methods for comparing uncertainty quantifications for material property predictions. Mach Learn Sci Technol. 2020;1(2):025006.

Wang T, Zhang C, Snoussi H. Machine learning approaches for thermoelectric materials research. Adv Funct Mater. 2020;30(5):1906041.
https://doi.org/10.1002/adfm.201906041.

Suzuki Y, Hino H, Hawai T, Saito K, Kotsugi M. Symmetry prediction and knowledge discovery from x-ray diffraction patterns using an interpretable machine learning approach. Sci Rep. 2020;10(1):21790.
https://doi.org/10.1038/s41598-020-77474-4.

Jablonka KM, Ongari D, Moosavi SM, Smit B. Big-data science in porous materials: Materials genomics and machine learning. Chem Rev. 2020;120(16):8066-8129.
https://doi.org/10.1021/acs.chemrev.0c00004.

Mortazavi B, Novikov IS, Podryabinkin EV. Exploring phononic properties of two-dimensional materials using machine learning interatomic potentials. Appl Mater Today. 2020;20:100685.
https://doi.org/10.1016/j.apmt.2020.100685.

Goodall REA, Lee AA. Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. Nat Commun. 2020;11(1):6280.
https://doi.org/10.1038/s41467-020-19964-7.

Bagherzadeh F, Shafighfard T. Ensemble machine learning approach for evaluating the material characterization of carbon nanotube-reinforced cementitious composites. Case Stud Constr Mater. 2022;17:e01126.
https://doi.org/10.1016/j.cscm.2022.e01126.

Zhang Z, Mansouri Tehrani A, Oliynyk AO, Day G, Brgoch J. Finding the next superhard material through ensemble learning. Adv Mater. 2021;33(10):2005112.
https://doi.org/10.1002/adma.202005112.

Mellit A, Kalogirou SA. Assessment of machine learning and ensemble methods for fault diagnosis of photovoltaic systems. Renew Energy. 2022;184:1074-90.
https://doi.org/10.1016/j.renene.2021.11.094.

Farooq F, Ahmed W, Akbar A, Aslam F. Predictive modeling for sustainable high-performance concrete from industrial wastes: A comparison and optimization of models using ensemble learners. J Clean Prod. 2021;292:126032.
https://doi.org/10.1016/j.jclepro.2021.126032.

Li QF, Song ZM. High-performance concrete strength prediction based on ensemble learning. Constr Build Mater. 2022;324:126694.
https://doi.org/10.1016/j.conbuildmat.2022.126694.

Sun Y, Li G, Zhang N, Chang Q, Xu J. Development of ensemble learning models to evaluate the strength of coal-grout materials. Int J Min Sci Technol. 2021;31(2):153-62.
https://doi.org/10.1016/j.ijmst.2021.01.003.

Feng DC, Liu ZT, Wang XD, Jiang ZM. Failure mode classification and bearing capacity prediction for reinforced concrete columns based on ensemble machine learning algorithm. Adv Eng Inform. 2020;45:101126.
https://doi.org/10.1016/j.aei.2020.101126.