Materials discovery

https://doi.org/10.1038/s42256-025-01061-3

Universal interatomic potentials shine in finding crystal structures

Ju Li

Various machine learning models have been developed in recent years for the discovery of crystal structures. Matbench Discovery, a new benchmark, offers an efficient way to identify the most promising architectures.

Determining the thermodynamic stability of atomic crystals with arbitrary assortment of different chemical elements is one of the most basic problems in materials science. Although real materials have trace impurities, defects, and metastable phases, which can make the actual structure and properties deviate from the ideal world, knowing the idealized crystalline phases is still often the first step. As calculating the ground-state atomic crystal structure from first-principles calculations is expensive, with a larger number of chemical elements and larger crystal primitive cells making the problem more difficult, researchers have resorted to machine learning (ML) methods. In this issue of *Nature Machine Intelligence*, Riebesell et al.[1] propose Matbench Discovery, a system for benchmarking machine intelligence in the task of discovering stable crystal structures (Fig. 1).

In the past few years, ML has had a key role in the discovery of all kinds of stable three-dimensional structures. As a notable example, Demis Hassabis and John Jumper were awarded one half of the Nobel Prize in Chemistry in 2024 for the development of AlphaFold, the ML-based program that predicts thermodynamically stable protein structures. This is analogous to ML-assisted crystal discovery in materials science[2]. Unsurprisingly, this field has also undergone explosive growth in the past decade, driven by developments in ML architectures, data generation, and automated training, testing and application workflows. On the architectural side, different types of neural networks that are sensitive to the crystal structure have blossomed, in addition to traditional ML methods such as random forests. On the data side, the open-access Materials Project (MP) database (established in 2011) provides more than 150,000 crystal structures whose potential energy is minimized by density functional theory (DFT) calculations, serving as a valuable starting point and reference.

Like the Critical Assessment of Structure Prediction (CASP) competition that eventually led to the Nobel Prize, the crystal structure community needs open benchmarks that enable fair and well-defined comparisons. Riebesell et al.[1] now propose Matbench Discovery, a benchmark system for just the ML architectures — that is, the training and testing data are fixed for all competing architectures — in discovering crystals with lower formation energy than the reference ones contained in the MP database. The results show that various universal interatomic potentials (UIPs)[3], the least restrictive form of structure-to-energy ML models, shine in the competition, taking all the top places on the leaderboard. The highest performing model, EquiformerV2 + DeNS, has a discovery acceleration factor (DAF) exceeding 5 — that is, it expedites the discovery of zero-kelvin thermodynamically stable structures at least five times faster than DFT calculations alone without ML. The results are verified by direct DFT calculations.

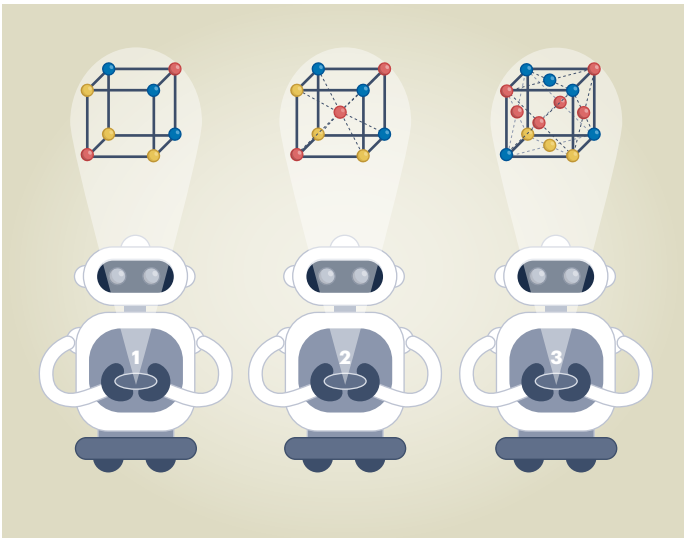

Fig. 1 | An evaluation framework for ML discovery of crystal structures. Rules need to be made to determine who is the champion in predicting thermodynamically stable crystal structures containing various chemical elements.

When considering these findings, some caveats apply. First, no experimental data were explicitly used in this benchmark. The Inorganic Crystal Structure Database (ICSD) has more than 290,000 experimental crystal structures resolved by X-ray diffraction, but they are not used for Matbench Discovery, nor are any new experimental synthesis or measurements carried out. Second, the ground truth in energy is taken to be DFT calculations. This can be problematic, as DFT calculations can have notable errors in predicting the formation energy, as large as 100 meV per atom (ref. 4). Third, the trial new crystals are sampled from the fixed WBM dataset[5], which come from chemical element substitutions of MP crystal structure motifs. New crystal-structure motifs that could potentially arise from randomized crystal structure prediction (CSP)[6] may be under-sampled. Fourth, vibrational contributions to the free energy are ignored, and one is replacing the free energy by 0 K potential energy. In short, Matbench Discovery considers "WBM crystals whose potential energy is lower than structures listed in the MP database according to 0 K DFT" as the ground truth for benchmarking the 14 competing ML architectures.

What is the value of such a benchmark system? The discovery of new crystalline phase beyond what experimental crystallographers

have found (even within the constrained permutation family of WBM methodology[5]) is one of the most important contributions to science from the ML community. I believe that this work is foundational to the field of materials science and engineering, as determining crystal structures at 0 K is the basis for most materials discoveries, both computational and experimental. The presented evidence that UIPs are best suited for this discovery task is definitive and very useful for the community. It is important to develop a fair benchmark to compare different methods and codes efficiently, which is an important service.

Given that UIPs are typically trained with much larger amounts of data, including far-from-equilibrium and saddle-point like configurations[3], the fact that UIP shines even when trained under limited near-equilibrium data is very encouraging. The future is bright for atomistic simulations using UIPs, as the map of thermodynamic equilibria has been illuminated.

Ju Li
Department of Nuclear Science and Engineering and Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA.
e-mail: liju@mit.edu

Published online: 26 June 2025

References

1. Riebesell, J. et al. *Nat. Mach. Intell.* https://doi.org/10.1038/s42256-025-01055-1 (2025).
2. Merchant, A. et al. *Nature* **624**, 80–85 (2023).
3. Takamoto, S., Okanohara, D., Li, Q.-J. & Li, J. *J. Mater.* **9**, 447–454 (2023).
4. Tang, H. et al. *Nat. Comput. Sci.* **5**, 144–154 (2205).
5. Wang, H.-C., Botti, S. & Marques, M. A. L. *npj Comput. Mater.* **7**, 12 (2021).
6. Glass, C. W., Oganov, A. R. & Hansen, N. *Comput. Phys. Commun.* **175**, 713 (2006).

Competing interests

The authors declare that they do not have competing interests.