Deep neural networks for accurate predictions of crystal stability

Weike Ye[1], Chi Chen[2], Zhenbin Wang[2], Iek-Heng Chu[2] & Shyue Ping Ong [2]

Predicting the stability of crystals is one of the central problems in materials science. Today, density functional theory (DFT) calculations remain comparatively expensive and scale poorly with system size. Here we show that deep neural networks utilizing just two descriptors—the Pauling electronegativity and ionic radii—can predict the DFT formation energies of $C_3A_2D_3O_{12}$ garnets and ABO_3 perovskites with low mean absolute errors (MAEs) of 7–10 meV atom^{-1} and 20–34 meV atom^{-1}, respectively, well within the limits of DFT accuracy. Further extension to mixed garnets and perovskites with little loss in accuracy can be achieved using a binary encoding scheme, addressing a critical gap in the extension of machine-learning models from fixed stoichiometry crystals to infinite universe of mixed-species crystals. Finally, we demonstrate the potential of these models to rapidly transverse vast chemical spaces to accurately identify stable compositions, accelerating the discovery of novel materials with potentially superior properties.

[1] Department of Chemistry and Biochemistry, University of California San Diego, 9500 Gilman Dr, Mail Code 0303, La Jolla, CA 92093-0448, USA.
[2] Department of NanoEngineering, University of California San Diego, 9500 Gilman Dr, Mail Code, 0448, La Jolla, CA 92093-0448, USA. Correspondence and requests for materials should be addressed to S.P.O. (email: ongsp@eng.ucsd.edu)

The formation energy of a crystal is a key metric of its stability and synthesizability. It is typically defined relative to constituent unary/binary phases (E_f) or the stable linear combination of competing phases in the phase diagram (E_{hull}, or energy above convex hull)[1]. In recent years, machine learning (ML) models trained on density functional theory (DFT)[2] calculations have garnered widespread interest as a means to scale quantitative predictions of materials properties[3–7], including energies of crystals. However, most previous efforts at predicting E_f or E_{hull} of crystals[5,8–12] using ML models have yielded mean absolute errors (MAEs) of 70–100 meV atom^{-1}, falling far short of the necessary accuracy for useful crystal stability predictions. This is because approximately 90% of the crystals in the Inorganic Crystal Structure Database (ICSD) have $E_{hull} < 70$ meV atom^{-1}[13], and the errors of DFT-calculated formation energies of ternary oxides from binary oxides relative to experiments are ~ 24 meV atom^{-1}[14].

We propose to approach the crystal stability prediction problem by using artificial neural networks (ANNs)[15], i.e., algorithms that are loosely modeled on the animal brain, to quantify well-established chemical intuition. The Pauling electronegativity and ionic radii guide much of our understanding about the bonding and stability of crystals today, for example, in the form of Pauling's five rules[16] and the Goldschmidt tolerance factor for perovskites[17]. Though these rules are qualitative in nature, their great success points to the potential existence of a direct relationship between crystal stability and these descriptors.

To probe these relationships, we choose, as our initial model system, the garnets, a large family of crystals with widespread technological applications such as luminescent materials for solid-state lighting[18] and lithium superionic conductors for rechargeable lithium-ion batteries[19,20]. Garnets have the general formula $C_3A_2D_3O_{12}$, where C, A and D denote the three cation sites with Wyckoff symbols 24c (dodecahedron), 16a (octahedron) and 24d (tetrahedron), respectively, in the prototypical cubic $Ia\overline{3}d$ garnet crystal shown in Fig. 1a. The distinct coordination environments of the three sites result in different minimum ionic radii ratios (and hence, species preference) according to Pauling's first rule. We further demonstrate the generalizability of our approach to the ABO_3 perovskites (Fig. 1b), another broad class of technologically important crystals[21–25].

In this work, we show that ANNs using only the Pauling electronegativity[26] and ionic radii[27] of the constituent species as the input descriptors can achieve extremely low MAEs of 7–10 meV atom^{-1} and 20–34 meV atom^{-1} in predicting the formation energies of garnets and perovskites, respectively. We also introduce two alternative approaches to extend such ANN models beyond simple unmixed crystals to the much larger universe of mixed cation crystals—a rigorously defined averaging scheme for the electronegativity and ionic radii for modeling complete cation disorder, and a novel binary encoding scheme to account for the effect of cation orderings with minimal increase in feature dimension. Finally, we demonstrate the application of the NN models in accurately and efficiently identifying stable compositions out of thousands of garnet and perovskite candidates, greatly expanding the space for the discovery of materials with potentially superior properties.

Results

Model construction and definitions. We start with the hypothesis that the formation energy E_f of a $C_3A_2D_3O_{12}$ garnet is some unknown function f of the Pauling electronegativities (χ) and Shannon ionic radii (r) of the species in the C,
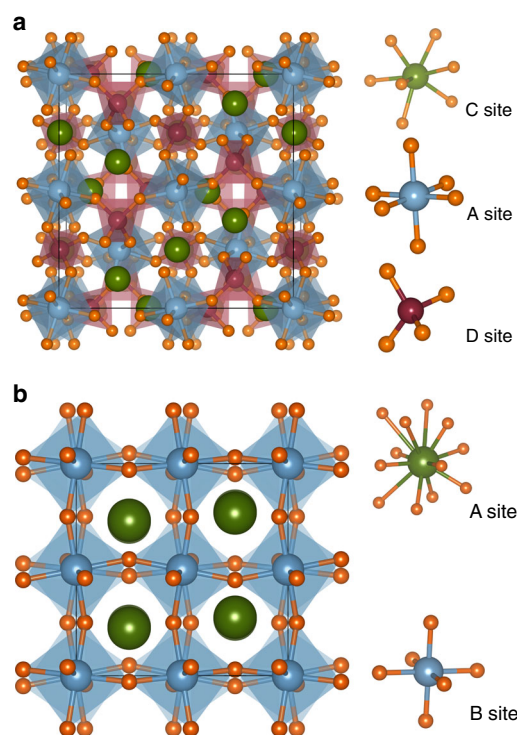

Fig. 1 Crystal structures of garnet and perovskite prototypes. **a** Crystal structure of $Ia\overline{3}d$ $C_3A_2D_3O_{12}$ garnet prototype. Green (C), blue (A), and red (D) spheres are atoms in the 24c (dodecahedron), 16a (octahedron), and 24d (tetrahedron) sites, respectively. The orange spheres are oxygen atoms. **b** Crystal structure of $Pnma$ ABO_3 perovskite prototype. Green (A) and blue (B) spheres are atoms in the 4c (cuboctahedron) and 4d (octahedron) sites, respectively. The orange spheres are oxygen atoms

A, and D sites, i.e.,

$$E_f = f(\chi_C, r_C, \chi_A, r_A, \chi_D, r_D) \tag{1}$$

Here, we define E_f as the change in energy in forming the garnet from binary oxides with elements in the same oxidation states, i.e., E_f^{oxide} as opposed to the more commonly used formation energy from the elements $E_f^{element}$ in previous works[8–11]. Using the $Ca_3Al_2Si_3O_{12}$ garnet (grossular) as an example, E_f^{oxide} is given by the energy of the reaction: $3CaO + Al_2O_3 + 3SiO_2 \rightarrow Ca_3Al_2Si_3O_{12}$. This choice of definition of E_f is motivated by two reasons. First, binary oxides are frequently used as synthesis precursors. Second, our definition ensures that garnets that share elements in the same oxidation states have E_f that are referenced to the same binary oxides, minimizing well-known DFT errors. In contrast, $E_f^{element}$ and E_{hull} are both poor target metrics for a ML model. $E_f^{element}$ suffers from non-systematic DFT errors associated with the incomplete cancellation of the self-interaction error in redox reactions[28], while E_{hull} is defined with respect to the linear combination of stable phases at the $C_3A_2D_3O_{12}$ composition in the C-A-D-O phase diagram, which can vary unpredictably even for highly similar chemistries. Henceforth, the notation E_f in this work refers to E_f^{oxide} unless otherwise stated. The binary oxides used to calculate the E_f for garnets and perovskites are listed in Supplementary Table 1 and 2, respectively.

Based on the universal approximation theorem[29], we may model the unknown function $f(\chi_C, r_C, \chi_A, r_A, \chi_D, r_D)$, which is clearly
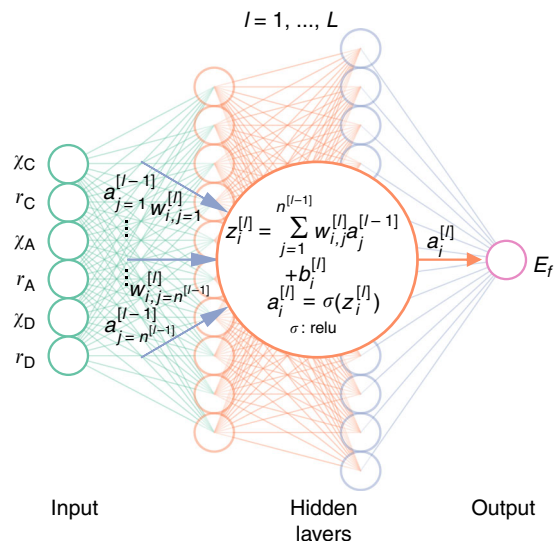
$$l = 1, ..., L$$

Fig. 2 General schematic of the artificial neural network. The artificial neural network (ANN) comprises an input layer of descriptors (the Pauling electronegativity and ionic radii on each site), followed by a number of hidden layers, and finally an output layer (E_f). The large circle in the centre shows how the output of the ith neuron in lth layer, $a_i^{[l]}$, is related to the received inputs from $(l-1)$th layer $a_j^{[l-1]}$. $w_{i,j}^{[l]}$ and $b_i^{[l]}$ denote the weight and bias between the jth neuron in $(l-1)$th layer and ith neuron in lth layer. σ is the activation function (rectified linear unit in this work). The ANN models were implemented using Keras[39] deep learning library with the Tensorflow[40] backend

non-linear (see Supplementary Fig. 1), using a feed-forward ANN, as depicted in Fig. 2. The loss function and evaluation metric are chosen to be the mean squared error (MSE) and MAE, respectively. We will denote the architecture of the ANN using $n^i - n^{[1]} - n^{[2]} - \cdots - 1$, where n^i and $n^{[l]}$ are the number of neurons in the input and lth hidden layer, respectively.

Neural network model for unmixed garnets. We developed an initial ANN model for unmixed garnets, i.e., garnets with only one type of species each in C, A, and D. A data set comprising 635 unmixed garnets was generated by performing full DFT relaxation and energy calculations (see Methods) on all charge-neutral combinations of allowed species (Supplementary Table 3) on the C, A, and D sites[30]. This dataset was randomly divided into training, validation, and test data in the ratio of 64:16:20. Using 50 repeated random sub-sampling cross validation, we find that a 6-24-1 ANN architecture yields a small root mean square error (RMSE) of 12 meV atom^{-1}, as well as the smallest standard deviation in the RMSE among the 50 sub-samples (Supplementary Fig. 2a). The training, validation and test MAEs for the optimized 6-24-1 model are ~7–10 meV atom^{-1} (Fig. 3a), an order of magnitude lower than the ~100 meV atom^{-1} achieved in previous ML models[5,8–10]. For comparison, the error in the DFT E_f of garnets relative to experimental values is around 14 meV atom^{-1} (Supplementary Table 4). Similar RMSEs are obtained for deep neural network (DNN) architectures containing two hidden layers (Supplementary Fig. 2b), indicating that a single-hidden-layer architecture is sufficient to model the relationship E_f and the descriptors.

Averaged neural network models for mixed garnets. To extend our model to mixed garnets, i.e., garnets with more than one type of species in the C, A, and D sites, we explored two alternative approaches—one based on averaging of descriptors, and another based on expanding the number of descriptors to account for the effect or species ordering. The data set for mixed garnets were

created using the same species pool, but allowing two species to occupy one of the sites. Mixing on the A sites was set at a 1:1 ratio, and that on the C and D sites was set at a 2:1 ratio, generating garnets of the form $C_3A'A''D_3O_{12}$ (211 compositions), $C'C''_2A_2D_3O_{12}$ (445 compositions), and $C_3A_2D'D''_2O_{12}$ (116 compositions). For each composition, we calculated the energies of all symmetrically distinct orderings within a single primitive unit cell of the garnet. All orderings must belong to a subgroup of the $Ia\bar{3}d$ garnet space group.

In the first approach, we characterized each C, A, or D site using weighted averages of the ionic radii and electronegativities of the species present in each site, given by the following expressions (see Methods):

$$r_{\text{avg}} = x r_X + (1 - x) r_Y \qquad (2)$$

$$\chi_{\text{avg}} = \chi_O - \sqrt{x(\chi_X - \chi_O)^2 + (1 - x)(\chi_Y - \chi_O)^2} \qquad (3)$$

where X and Y are the species present in a site with fraction x and $(1-x)$, respectively, and O refers to the element oxygen. The implicit assumption in this "averaged" ANN model is that species X and Y are completely disordered, i.e., different orderings of X and Y result in negligible DFT energy differences.

Using the same 6-24-1 ANN architecture, we fitted an "averaged" model using the energy of the ground state ordering of the 635 unmixed and 772 mixed garnets. We find that the training, validation, and test MAEs of the optimized model are 22, 26, and 26 meV atom^{-1}, respectively (Supplementary Fig. 3a). These MAEs are about double that of the unmixed ANN model, but still comparable to the error of the DFT E_f relative to experiments. The larger MAEs may be attributed to the fact that the effect of species orderings on the crystal energy is not accounted for in this "averaged" model.

Ordered neural network model for mixed garnets. In the second approach, we undertook a more ambitious effort to account for the effect of species orderings on crystal energy. Here, we discuss the results for species mixing on the C site only, for which the largest number of computed compositions and orderings is available. For 2:1 mixing, there are 20 symmetrically distinct orderings within the primitive garnet cell, which can be encoded using a 5-bit binary array $[b_0, b_1, b_2, b_3, b_4]$. This binary encoding scheme is significantly more compact that the commonly used one-hot encoding scheme, and hence, minimizes the increase in the descriptor dimensionality. We may then modify Eq. 1 as follows:

$$E_f = f(\chi_{C'}, r_{C'}, \chi_{C''}, r_{C''}, \chi_A, r_A, \chi_D, r_D, b_0, b_1, b_2, b_3, b_4) \qquad (4)$$

where the electronegativities and ionic radii of both species on the C sites are explicitly represented. In contrast to the "averaged" model, we now treat the 20 ordering-E_f pairs at each composition as distinct data points. Each unmixed composition was also included as 20 data points with the same descriptor values and E_f, but different binary encodings.

We find that a two-hidden-layer DNN is necessary to model this more complex composition-ordering-energy relationship. The final optimized 13-22-8-1 model exhibits overall training, validation and test MAEs of ~11–12 meV atom^{-1} on the entire unmixed and mixed dataset (Supplementary Fig. 3b). The comparable MAEs between this extended DNN model and the unmixed ANN model is clear evidence that the DNN model has successfully captured the additional effect of orderings on E_f. We note that the average standard deviation of the predicted E_f of
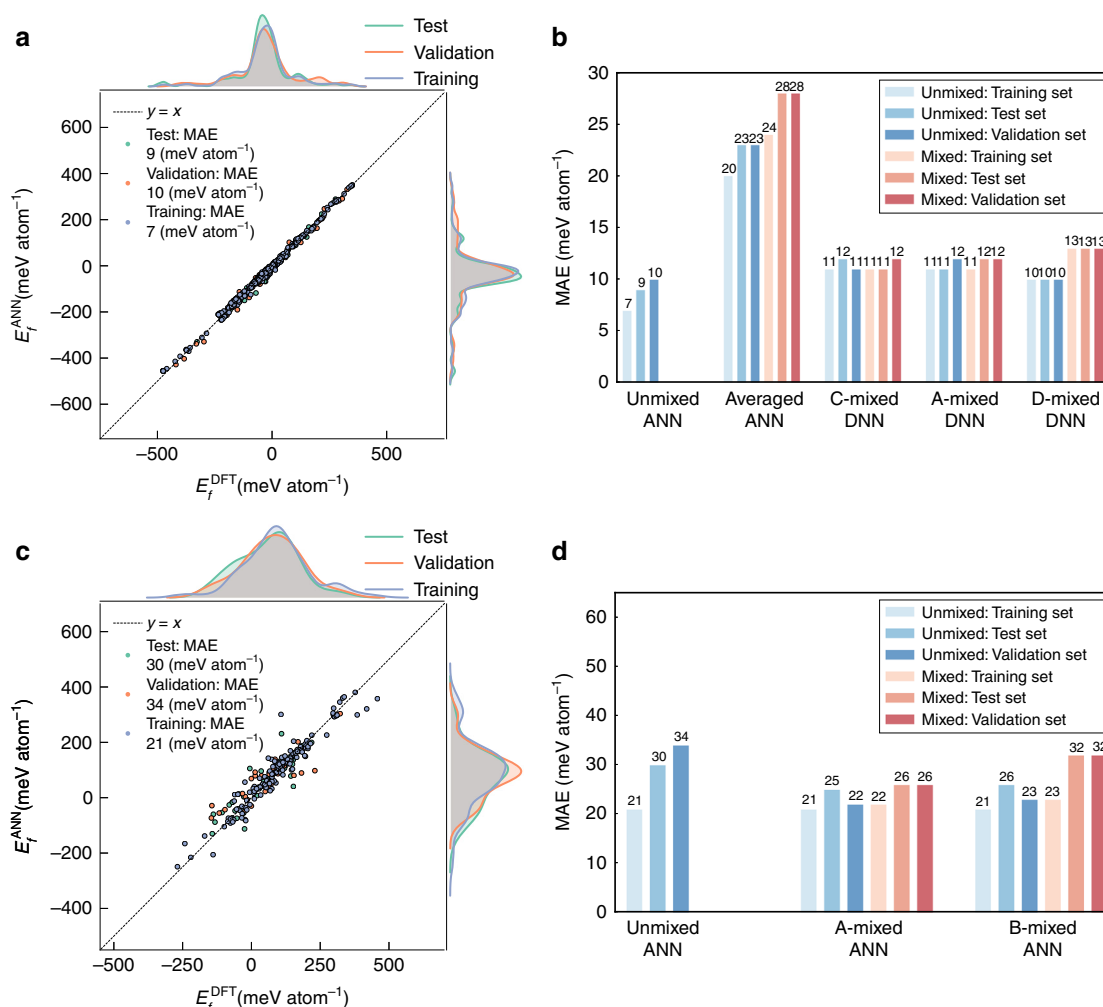
Fig. 3 Performance of artificial neural network (ANN) models. **a** Plot of E_f^{ANN} against E_f^{DFT} of unmixed garnets for optimized 6-24-1 ANN model. The histograms at the top and right show that the training, validation and test sets contain a good spread of data across the entire energy range of interest with standard deviations of 122–134 meV atom^{-1}. Low mean absolute errors (MAEs) in E_f of 7, 10, and 9 meV atom^{-1} are observed for the training, validation, and test sets, respectively. **b** MAEs in E_f of unmixed and mixed samples in training, validation, and test sets of all garnet models. The C-, A- and D-mixed deep neural networks (DNNs) have similar MAEs as the unmixed ANN model, indicating that the neural network has learned the effect of orderings on E_f. Each C-, A- and D-mixed composition has 20, 18, and 7 distinct orderings, respectively, which are encoded using 5-bit, 5-bit, and 3-bit binary arrays, respectively. **c** MAEs in E_f of unmixed and mixed samples for training, validation and test sets of unmixed perovskites for 4-12-1 ANN model. The E_f^{DFT} of training, validation, and test sets similarly contain a good spread of data across the entire energy range of interest with standard deviations of 104–122 meV atom^{-1}. Low mean absolute errors (MAEs) in E_f of 21, 34, and 30 meV atom^{-1} are observed for the training, validation, and test sets, respectively. **d** MAEs in E_f for training, validation, and test sets of all perovskite models. Each A- and B- mixed perovskite compositions has ten distinct orderings, which are both encoded using 4-bit binary arrays. The black lines (dashed) in (**a**, **c**) are the identity lines serving as references

different orderings of unmixed compositions using this extended DNN model is only 2.8 meV atom^{-1}, indicating that the DNN has also learned the fact that orderings of the same species on a particular site have little effect on the energy. Finally, similar MAEs can be achieved for A and D site mixing (Supplementary Fig. 3c and 3d) using the same approach.

Stability classification of garnets using ANN models. While E_f is a good target metric for a predictive ANN model, the stability of a crystal is ultimately characterized by its E_{hull}. Using the predicted E_f from our DNN models and pre-calculated DFT data from the Materials Project[31], we have computed E_{hull} by constructing the 0 K C-A-D-O phase diagrams. From Fig. 4a, we may observe that the extended C-mixed DNN model can achieve a >90% accuracy in classifying stable/unstable unmixed garnets at a strict E_{hull} threshold of 0 meV atom^{-1} and rises rapidly with increasing threshold. Similarly, high classification accuracies of

greater than 90% are achieved for all three types of mixed garnets. Given the great flexibility of the garnet prototype in accommodating different species, there are potentially millions of undiscovered compositions. Even using our restrictive protocol of single-site mixing in specified ratios, 8427 mixed garnet compositions can be generated, of which 2307 are predicted to have E_{hull} of 0 meV atom^{-1}, i.e., potentially synthesizable (Supplementary Fig. 4a). A web application that computes E_f and E_{hull} for any garnet composition using the optimized DNNs has been made publicly available for researchers at http://crystals.ai.

Neural network models for unmixed and mixed perovskites. To demonstrate that our proposed approach is generalizable and not specific to the garnet crystal prototype, we have constructed similar neural network models using a dataset of 240 unmixed, 222 A-mixed and 80 B-mixed ABO_3 perovskites generated using the species in Supplementary Table 5. We find that a 4-12-1
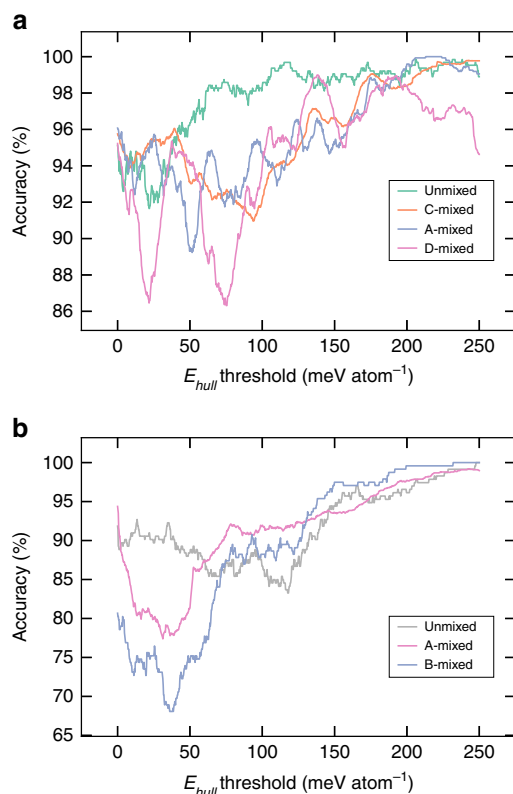
Fig. 4 Accuracy of stability classification. Plots of the accuracy of stability classification of the ANN models compared to DFT as a function of the E_{hull} threshold for **a**. garnets, and **b**. perovskites. The accuracy is defined as the sum of the true positive and true negative classification rates. A true positive (negative) means that the E_{hull} for a particular composition predicted from the optimized artificial neural network model and DFT are both below (above) the threshold. For the mixed compositions, an E_{hull} is calculated for all orderings (20, 7, and 18 orderings per composition for C-, A-, and D-mixed garnets, respectively, and ten orderings per composition for both A- and B-mixed perovskites)

single-hidden-layer neural network is able to achieve MAEs of 21–34 meV atom^{-1} in the predicted E_f for unmixed perovskites (Fig. 3c), while two 10-24-1 neural networks are able to achieve MAEs of 22–39 meV atom^{-1} in the E_f of the mixed perovskites (Supplementary Fig. 5). These MAEs are far lower than those of prior ML models of unmixed perovskites, which generally have MAEs of close to 100 meV atom^{-1} or higher[9,16]. As shown in Fig. 3b, the accuracy of classifying stable versus unstable perovskites exceeds 80% at a strict E_{hull} threshold of 0 meV atom^{-1} and maintains at above 70% at a loosened E_{hull} threshold of 30 meV atom^{-1}. During the review of this work, a new work by Li et al.[32] reported achieving comparable MAEs of ~28 meV atom^{-1} in predicting the E_{hull} of perovskites using a kernel ridge regression model. However, this performance was achieved using a set of 70 descriptors, with model performance sharply dropping with less than 70 descriptors. Furthermore, Li et al.'s model is restricted to perovskites with $E_{hull} < 400$ meV atom^{-1} and only a single ordering for each mixed perovskite, while in this work, the highest E_{hull} is 747 meV atom^{-1} for the perovskite dataset and all symmetrically distinct orderings on the A and B sites within a √2×√2×1 orthorhombic conventional perovskite unit cell (ten structures each) are considered.

Discussion

To summarize, we have shown that NN models can quantify the relationship between traditionally chemically intuitive descriptors,

such as the Pauling electronegativity and ionic radii, and the energy of a given crystal prototype. A key advantage of our proposed NN models is that they rely only on an extremely small number (two) of site-based descriptors, i.e., no structural degrees of freedom are considered beyond the ionic radii of a particular species in a site and the ordering of the cations in the mixed oxides. This is in stark contrast to most machine-learning models in the literature utilizing a large number of correlated descriptors, which render such models highly susceptible to overfitting, or machine-learning force-fields, which can incorporate structural and atomic degrees of freedom but at a significant loss of transferability to different compositions. Most importantly, we derive two alternative approaches—a rigorously defined averaging scheme to model complete cation disorder and a binary encoding scheme to account for the effect of orderings—to extend high-performing unmixed deep learning models to mixed cation crystals with little/no loss in error performance and minimal increase in descriptor dimensionality. It should be noted that our NN models are still restricted to the garnet and perovskite compositions (with or without cation mixing) with no vacancies, though further extensions to other common crystal structure prototypes and to account for vacancies should in principle be possible. Finally, we show how predictive models of E_f can be combined with existing large public databases of DFT computed energies to predict E_{hull} and hence, phase stability. These capabilities can be used to efficiently traverse large chemical spaces of unmixed and mixed crystals to identify stable compositions and orderings, greatly accelerating the potential for novel materials discovery.

Methods

DFT calculations. All DFT calculations were performed using Vienna ab initio simulation package (VASP) within the projector augmented-wave approach[33,34]. Calculation parameters were chosen to be consistent with those used in the Materials Project, an open database of pre-computed energies for all known inorganic materials[31]. The Perdew-Burke-Ernzehof generalized gradient approximation exchange-correlation functional[35] and a plane-wave energy cut-off of 520 eV were used. Energies were converged to within 5×10^{-5} eV atom^{-1}, and all structures were fully relaxed. For mixed compositions, symmetrically distinct orderings within the 80-atom primitive garnet unit cell and the 40-atom √2×√2×1 orthorhombic perovskite supercell were generated using the enumlib library[36] via the Python Materials Genomics package.[37]

Training of ANNs. Training of the ANNs was carried out using the Adam optimizer[38] at a learning rate of 0.2, with the mean square error of E_f as the loss metric. For each architecture, we ran with a random 64:16:20 split of training, validation and test data, i.e., random sub-sampling cross validation.

Electronegativity averaging. Pauling's definition of electronegativity is based on an "additional stabilization" of a heteronuclear bond X–O compared to average of X–X and O–O bonds, as follows.

$$(\chi_X - \chi_O)^2 = E_d(XO) - \frac{E_d(XX) + E_d(OO)}{2}$$

where χ_X and χ_O are the electronegativities of species X and O, respectively, and E_d is the dissociation energy of the bond in parentheses. Here, O refers to oxygen.

For a disordered site containing species X and Y in the fractions x and $(1-x)$, respectively, we obtain the following:

$$\left(\chi_{X_x Y_{1-x}} - \chi_O\right)^2 = x E_d(XO) + (1-x) E_d(YO)$$
$$- \frac{x E_d(XX) + (1-x) E_d(YY) + E_d(OO)}{2}$$
$$= x(\chi_X - \chi_O)^2 + (1-x)(\chi_Y - X_O)^2$$

We then obtain the effective electronegativity for the disordered site as follows:

$$\chi_{X_x Y_{1-x}} = \chi_O - \sqrt{x(\chi_X - \chi_O)^2 + (1-x)(\chi_Y - \chi_O)^2}$$

Data availability

The datasets generated during and/or analysed during the current study are available in the GitHub repository https://github.com/materialsvirtuallab/garnetdnn as well as the Dryad Digital Repository (doi: 10.5061/dryad.760r5b6). A web application that estimates E_f and E_{hull} for any given garnet or perovskite composition using the optimized DNNs is available at http://crystals.ai/.

References

1. Ong, S. P., Wang, L., Kang, B. & Ceder, G. Li−Fe−P−O2 phase diagram from first principles calculations. *Chem. Mater.* **20**, 1798–1807 (2008).
2. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864 (1964).
3. Pilania, G., Wang, C., Jiang, X., Rajasekaran, S. & Ramprasad, R. Accelerating materials property predictions using machine learning. *Sci. Rep.* **3**, 2810 (2013).
4. Lee, J., Seko, A., Shitara, K., Nakayama, K. & Tanaka, I. Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques. *Phys. Rev. B* **93**, 115104 (2016).
5. Schmidt, J. et al. Predicting the thermodynamic stability of solids combining density functional theory and machine learning. *Chem. Mater.* **29**, 5090–5103 (2017).
6. Pilania, G. et al. Machine learning bandgaps of double perovskites. *Sci. Rep.* **6**, 19375 (2016).
7. Isayev, O. et al. Universal fragment descriptors for predicting properties of inorganic crystals. *Nat. Commun.* **8**, 15679 (2017).
8. Meredig, B. et al. Combinatorial screening for new materials in unconstrained composition space with machine learning. *Phys. Rev. B* **89**, 094104 (2014).
9. Faber, F. A., Lindmaa, A., Von Lilienfeld, O. A. & Armiento, R. Machine learning energies of 2 million elpasolite (ABC_2D_6). *Phys. Rev. Lett.* **117**, 135502 (2016).
10. Ward, L., Agrawal, A., Choudhary, A. & Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Comput. Mater.* **2**, 16028 (2016).
11. Ward, L. et al. Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations. *Phys. Rev. B* **96**, 024104 (2017).
12. Seko, A., Hayashi, H., Nakayama, K., Takahashi, A. & Tanaka, I. Representation of compounds for machine-learning prediction of physical properties. *Phys. Rev. B* **95**, 144110 (2017).
13. Sun, W. et al. The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225–e1600225 (2016).
14. Hautier, G., Ong, S. P., Jain, A., Moore, C. J. & Ceder, G. Accuracy of density functional theory in predicting formation energies of ternary oxides from binary oxides and its implication on phase stability. *Phys. Rev. B - Condens. Matter Mater. Phys.* **85**, 155208 (2012).
15. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
16. Pauling, L. The principles determining the structure of complex ionic crystals. *J. Am. Chem. Soc.* **51**, 1010–1026 (1929).
17. Goldschmidt, V. M. Die Gesetze der Krystallochemie. *Naturwissenschaften* **14**, 477–485 (1926).
18. Nakamura, S. Present performance of InGaN-based blue/green/yellow LEDs. *Proc. SPIE* 3002, 26–35 (1997).
19. O'Callaghan, M. P., Lynham, D. R., Cussen, E. J. & Chen, G. Z. Structure and ionic-transport properties of lithium-containing garnets $Li_3Ln_3Te_2O_{12}$ (Ln = Y, Pr, Nd, Sm−Lu). *Chem. Mater.* **18**, 4681–4689 (2006).
20. Peng, H., Wu, Q. & Xiao, L. Low temperature synthesis of $Li_5La_3Nb_2O_{12}$ with cubic garnet-type structure by sol–gel process. *J. Sol.-Gel Sci. Technol.* **66**, 175–179 (2013).
21. Kobayashi, K.-I., Kimura, T., Sawada, H., Terakura, K. & Tokura, Y. Room-temperature magnetoresistance in an oxide material with an ordered double-perovskite structure. *Nature* **395**, 677–680 (1998).
22. Cava, R. J. et al. Bulk superconductivity at 91 K in single-phase oxygen-deficient perovskite $Ba_2YCu_3O_{9 − \delta}$. *Phys. Rev. Lett.* **58**, 1676–1679 (1987).
23. Cohen, R. E. Origin of ferroelectricity in perovskite oxides. *Nature* **358**, 136–138 (1992).
24. Grinberg, I. et al. Perovskite oxides for visible-light-absorbing ferroelectric and photovoltaic materials. *Nature* **503**, 509–512 (2013).
25. Green, M. A., Ho-Baillie, A. & Snaith, H. J. The emergence of perovskite solar cells. *Nat. Photonics* **8**, 506–514 (2014).
26. Pauling, L. The nature of the chemical bond. IV. The energy of single bonds and the relative electronegativity of atoms. *J. Am. Chem. Soc.* **54**, 3570–3582 (1932).
27. Shannon, R. D. Revised effective ionic radii and systematic studies of interatomie distances in halides and chaleogenides. *Acta Cryst.* **A32**, 751–767 (1976).
28. Wang, L., Maxisch, T. & Ceder, G. Oxidation energies of transition metal oxides within the GGA + U framework. *Phys. Rev. B* **73**, 195107 (2006).
29. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **4**, 251–257 (1991).
30. Granat-struktur, D., Übersicht, E., Kationen, D. & Ionenverteilung, D. Crystal chemistry of the garnet. *Z. für Krist. - Cryst. Mater.* **47**, 1989 (1999).
31. Jain, A. et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
32. Li, W., Jacobs, R. & Morgan, D. Predicting the thermodynamic stability of perovskite oxides using machine learning models. *Comput. Mater. Sci.* **150**, 454–463 (2018).
33. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
34. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
35. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
36. Hart, G. L. W., Nelson, L. J. & Forcade, R. W. Generating derivative structures at a fixed concentration. *Comput. Mater. Sci.* **59**, 101–107 (2012).
37. Ong, S. P. et al. Python materials genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
38. Kingma, D. P. & Jimmy Ba, Adam: a method for stochastic optimization. Preprint at https://arxiv.org/pdf/1412.6980 (2016).
39. Chollet, F. et al. Keras. http://keras.io (2015).
40. Abadi, M. M. et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. http://www.tensorflow.org/ (2015).

Author contributions

S.P.O., W.Y. and C.C. proposed the concept. W.Y. carried out the calculations and analysis with the help from C.C., Z.W. and I.C. W.Y. prepared the initial draft of the manuscript. All authors contributed to the discussions and revisions of the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06322-x.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.