







## ARTICLE OPEN



## Closed-loop superconducting materials discovery

Elizabeth A. Pogue<sup>1,6</sup>, Alexander New<sup>1,6</sup>, Kyle McElroy<sup>1</sup>, Nam Q. Le<sup>1</sup>, Michael J. Pekala<sup>1</sup>, Ian McCue<sup>1,2</sup>, Eddie Gienger<sup>1</sup>, Janna Domenico<sup>1</sup>, Elizabeth Hedrick<sup>3</sup>, Tyrel M. McQueen<sup>1,3,4,5</sup>, Brandon Wilfong<sup>4,5</sup>, Christine D. Piatko<sup>1</sup>, Christopher R. Ratto<sup>1</sup>, Andrew Lennon<sup>1</sup>, Christine Chung<sup>1</sup>, Timothy Montalbano<sup>1</sup>, Gregory Bassen<sup>4,5</sup> and Christopher D. Stiles<sup>1</sup>

Discovery of novel materials is slow but necessary for societal progress. Here, we demonstrate a closed-loop machine learning (ML) approach to rapidly explore a large materials search space, accelerating the intentional discovery of superconducting compounds. By experimentally validating the results of the ML-generated superconductivity predictions and feeding those data back into the ML model to refine, we demonstrate that success rates for superconductor discovery can be more than doubled. Through four closed-loop cycles, we report discovery of a superconductor in the Zr-In-Ni system, re-discovery of five superconductors unknown in the training datasets, and identification of two additional phase diagrams of interest for new superconducting materials. Our work demonstrates the critical role experimental feedback provides in ML-driven discovery, and provides a blueprint for how to accelerate materials progress.

*npj Computational Materials* (2023)9:181; <https://doi.org/10.1038/s41524-023-01131-3>

## INTRODUCTION

The discovery of novel materials drives industrial innovation<sup>1–3</sup>, although the pace of discovery tends to be slow due to the infrequency of “Eureka!” moments<sup>4,5</sup>. These moments are typically tangential to the original target of the experimental work: “accidental discoveries”. Here we demonstrate the acceleration of intentional materials discovery—targeting material properties of interest while generalizing the search to a large materials space with machine learning (ML) methods combined with experiment in a feedback loop. We demonstrate a closed-loop joint ML-experimental discovery process targeting unreported superconducting materials, which have industrial applications ranging from quantum computing to sensors to power delivery<sup>6–9</sup>. By closing the loop, i.e., by experimentally testing the results of the ML-generated superconductivity predictions and feeding data back into the ML model to refine, we demonstrate that success rates for superconductor discovery can be more than doubled<sup>10</sup>. In four closed-loop cycles, we discovered an unreported superconductor in the Zr-In-Ni system, re-discovered five superconductors unknown in the training datasets, and identified two additional phase diagrams of interest for superconducting materials. Our work demonstrates the critical role experimental feedback provides in ML-driven discovery, and provides definite evidence that such technologies can accelerate discovery even in the absence of knowledge of the underlying physics.

Statistical approaches have long aimed to better understand and predict superconductivity<sup>11</sup>, most recently through the use of black-box ML methods<sup>12–18</sup>. Although resulting in numerous predictions, these studies have not yielded previously unreported families of superconductors, likely not only because of difficulties in extrapolating beyond known families, but also because the predicted materials have chemical attributes that make them unlikely to be superconducting—whether it is highly localized chemical bonding, e.g., those containing polyatomic anions, or an extreme metastability that precludes synthesizability. Further, existing works have treated materials and databases of material

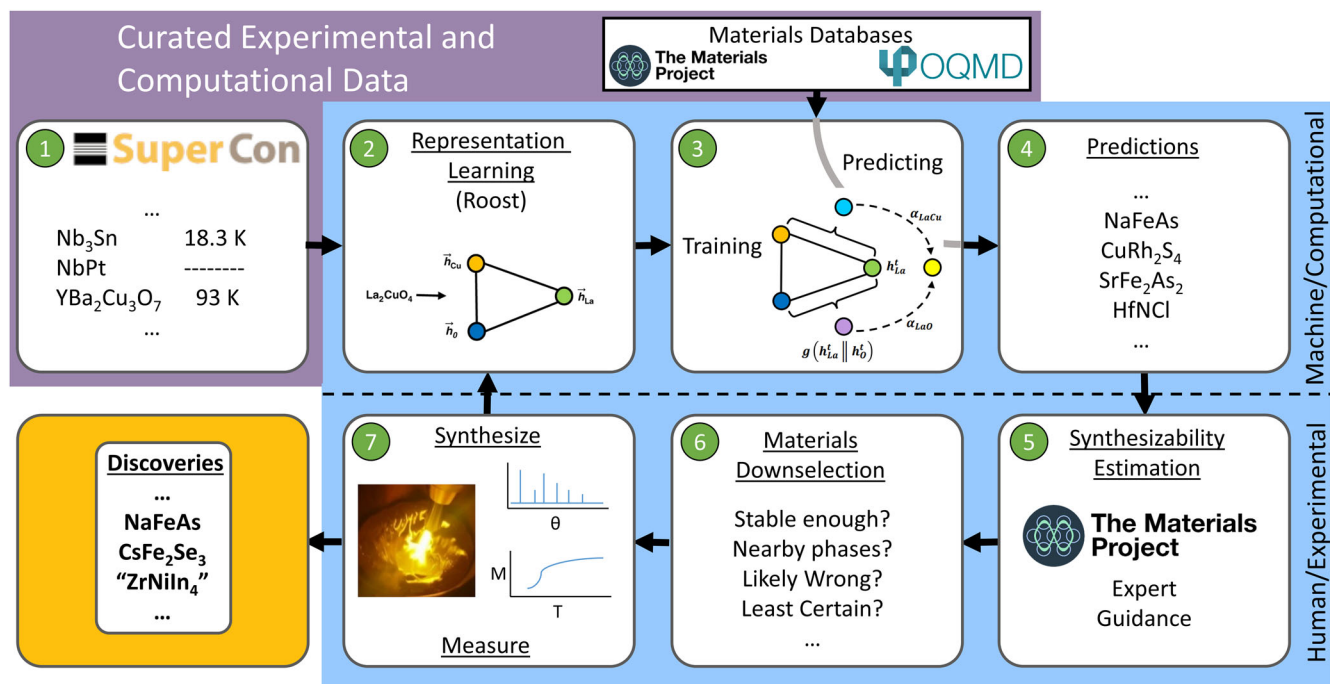
properties as fixed snapshots rather than evolving systems, which limits the ability of ML models to learn over sparse data.

Here we report on combining ML techniques with materials science and physics expertise to “close the loop” of materials discovery (Fig. 1). We demonstrate how to make ML models generalize across diverse materials spaces, to identify superconductors that are dissimilar to ones in the training corpus. By alternating between ML property prediction and experimental verification, we are able to systematically improve the fidelity of ML property prediction in regimes sparsely represented by existing materials databases. Crucially, this adds both negative data (materials incorrectly predicted to be superconductors) and positive data (materials correctly predicted) to ML training, enabling the ML model’s overall representation of the space of materials to be iteratively refined. The result is a ML model for predicting superconductivity that doubles the rate of successful predictions<sup>10</sup>, demonstrating the acceleration of materials discovery by combining human and machine insight.

Our process uses active learning<sup>19</sup> to iteratively select data points to be added to a training set. In particular, we select materials that are both predicted to be high transition temperature ( $T_c$ ) superconductors and are sufficiently distinct from known superconductors. We also leverage human domain expertise to further refine selections. When the predictive model incorrectly predicts non-superconductors as superconductors, this valuable negative data helps refine the model’s prediction surface.

A key attribute of our work is that the training data used in the ML models is not static, but evolves as the closed-loop process proceeds. A ML model that is employing a closed loop framework, actively sampling regions of previously unexplored spaces of materials, and continually acquiring new data cannot have a concise picture of convergence, and it is changed with every loop. Thus instead of a traditional convergence metric (e.g., looking for a flattening of loss versus number of training epochs for a convolutional neural network), we leverage goal-based metrics—when the model successfully predicts superconductors not in the training set or the human in the loop assesses that

<sup>1</sup>Research and Exploratory Development Department, Johns Hopkins University Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel 20723 MD, USA. <sup>2</sup>Department of Materials Science and Engineering, Northwestern University, 2220 Campus Drive, Evanston 60208 IL, USA. <sup>3</sup>Department of Materials Science and Engineering, Johns Hopkins University, 3400 N. Charles Street, Baltimore 21218 MD, USA. <sup>4</sup>Department of Chemistry, Johns Hopkins University, 3400 N. Charles Street, Baltimore 21218 MD, USA. <sup>5</sup>Institute for Quantum Matter, William H. Miller III Department of Physics and Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore 21218 MD, USA. <sup>6</sup>These authors contributed equally: Elizabeth A. Pogue, Alexander New. ✉email: [Lisa.Pogue@jhuapl.edu](mailto:Lisa.Pogue@jhuapl.edu); [mcqueen@jhu.edu](mailto:mcqueen@jhu.edu); [Christopher.Stiles@jhuapl.edu](mailto:Christopher.Stiles@jhuapl.edu)



**Fig. 1 The closed loop discovery process.** Starting from curated experimental data of known superconductors (1), compositional information is first transformed into a representation suitable for learning using the RooSt<sup>27</sup> framework (2). After initial training of the ML model (3), we provide new compositions not known to the ML model from other sources, and obtain predictions of superconducting behavior (4). The synthesizability of these predictions is assessed using a combination of computational thermodynamic data and expert insight (5). Materials downselection (6) occurs with human input based on multiple criteria to maximize the impact experimental work has on model improvement and related factors. Chosen materials are then synthesized and structural and physical properties measured (7). Results are then fed back into the learning process, in addition to generating discoveries. Further details on the closed loop process are in “Closed-loop discovery process” in Methods. RooSt images used with permission, CC-BY-4.0 license<sup>27</sup>.

**Table 1.** Superconductors rediscovered by machine learning.

Compound	Iteration	Database and Comments	$T_c$
NaFeAs	2	Materials Project, reduced topotactically in water	25 K <sup>53</sup>
CuRh <sub>2</sub> S <sub>4</sub>	2	Materials Project	4.7 K <sup>54</sup>
SrFe <sub>2</sub> As <sub>2</sub>	2	Materials Project, reduced topotactically in water	25 K <sup>55</sup>
HfNCl	3 and 4	Materials Project, OQMD, intercalated by lithium	40 K <sup>56</sup>
CsFe <sub>2</sub> Se <sub>3</sub>	3	Analogs CsFe <sub>4-x</sub> Se <sub>4</sub> (5 K, > ~34.55 GPa) <sup>57</sup> , BaFe <sub>2</sub> Se <sub>3</sub> (11 K, > ~12.7 GPa) <sup>58</sup> , and BaFe <sub>2</sub> S <sub>3</sub> (24 K, > ~10 GPa) <sup>59</sup>	

Since the model only took stoichiometry into account, predictions where small stoichiometry changes led to superconductivity were considered successful discoveries.

model outputs are sufficiently distinct and chemically plausible from prior predictions. This helps avoid model overfitting by terminating the process earlier than a traditional metric, while maximizing the usefulness of the new experimental data to further refine the model.

Utilizing this iterative “closed-loop” approach, we rediscover five known superconductors outside of the ML model’s training set, Table 1. These materials come from a wide variety of families: iron pnictides, doped 2D ternary transition metal nitride halides, and intermetallics, Table 2. We then further report the

**Table 2.** Distribution of  $T_c$  values in SuperCon.

	$T_c < 2$	$2 \leq T_c < 20$	$20 \leq T_c$
# Materials	5802	6182	4320

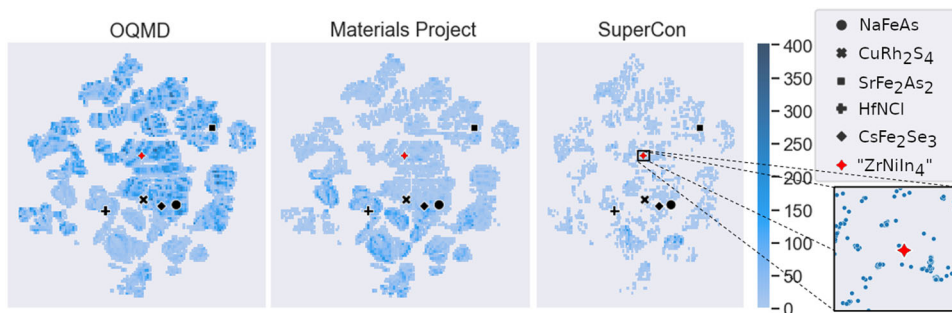
Distribution of SuperCon entries by tertile in our classification formulation. We treat the prediction problem as a three-class classification problem.

discovery of a previously unreported superconductor in the Zr-In-Ni phase diagram, and identified two other phase diagrams of interest (Zr-In-Cu and Zr-Fe-Sn).

## RESULTS AND DISCUSSION

### Model generation

For the initial prediction step of the closed-loop approach, we trained an ML model to predict the superconducting transition temperature,  $T_c$ , of candidate materials. Our primary source of training data, SuperCon<sup>20</sup>, contains compositions of known superconductors. Only the materials’ compositions were used to train the ML model for predicting  $T_c$  since SuperCon did not contain additional structural information. Materials Project (MP)<sup>21</sup> and Open Quantum Materials Database (OQMD)<sup>22</sup>, some of the largest public sets of computational materials data, supplied candidate compositions to be screened for superconductivity. These two databases do not contain any  $T_c$  data. These three datasets are visualized in Fig. 2 using a joint representation. Crucially, the amount of data for which we have superconducting information is much smaller than our other sources of data and is not uniformly sampled across the joint space.



**Fig. 2 Training data sparseness and finding non-derivative superconductors.** Histograms of the concentration of materials from a Uniform Manifold and Projection (UMAP)<sup>60</sup> embedding of OQMD (without superconductivity information), MP (without superconductivity information), and SuperCon (superconductivity information), based on Magpie<sup>30</sup> descriptors for the datasets. The embedding is learned from concatenation of Magpie descriptors obtained from all three datasets; the same axis limits are used across each subplot. These maps show the sparseness of knowledge of data about superconductivity compared to that of all known and predicted compounds in these open databases.  $T_c$  is not part of the Magpie descriptors and, therefore, did not influence the representation. The five black symbols indicate rediscovered superconductors (Table 1), and the red symbol our superconductor, near “ZrNiIn<sub>4</sub>”. The inset on the right highlights the local region in which “ZrNiIn<sub>4</sub>” is found, which is sparse and far from the known and rediscovered superconductors.

It is well-known<sup>23</sup> that when ML methods make predictions on data outside of their training data distribution, accuracy often suffers; this is often called the *out-of-distribution generalization problem*. In cheminformatics<sup>24</sup>, it is common to assess whether a dataset is within the distribution of a training dataset by seeing how far, in some representative metric space, its points are from the training dataset: as the difference between the distribution of new data and the training data increases, the likelihood that a model will accurately predict their properties decreases. To improve assessment of generalization, out-of-distribution data may be simulated by creating validation sets that split based on non-random criteria like Murcko scaffold<sup>25</sup> or cluster identity, the latter being the leave-one-cluster-out cross-validation (LOCO-CV) strategy<sup>26</sup>.

In “Model Validation” in Methods, we apply LOCO-CV in a simulated superconductor-identification problem. We show that, although a strong ML model is capable of fitting the training set well and generalizing to out-of-distribution test data, it fails to make accurate predictions of superconducting status on out-of-distribution data. Because existing superconductor datasets are not sufficient to enable accurate identification of unreported superconductors, this motivates the need for multiple iterations of model training, candidate selection, candidate synthesis, and model retraining.

We rely on a recent ML model for chemical property prediction, Representation learning from Stoichiometry (RooSt)<sup>27</sup> (see “Computational Methods and Uncertainty” in Methods and the SI), to predict a material’s superconductivity using only its stoichiometry (i.e., ignoring the material’s crystal structure). Although not as immediately powerful as approaches incorporating structural information<sup>16,28,29</sup>, it enables greater predictive sensitivity because materials compositions can be tested without knowledge of the structure.

### Superconductivity-specific considerations

After training an ensemble of RooSt models using the SuperCon database, we apply them to our set of potential superconductors (i.e., MP and OQMD). We filter for materials likely to be high- $T_c$  superconductors, and then selected materials are synthesized and characterized, enabling the ML model to be retrained in further loop iterations.

A risk of searching for superconductors from a static list of candidates is that while a material in MP or OQMD may not have the exact composition as a superconductor, it may have a composition extremely close in terms of stoichiometry, such as MgB<sub>2</sub> vs. Mg<sub>33</sub>B<sub>67</sub>. Thus, every time we produce a new list of candidates, we identify each candidate’s minimal Euclidean distance, in Magpie-space<sup>30</sup>, to any point in our training data, and we remove candidates too close to SuperCon.

It is not practical to experimentally verify all ML predictions. The costs associated with fabricating and characterizing a new material are high; hence we are only able to experimentally analyze a small subset of the ML predictions.

The MP and OQMD databases both contain calculated stability information not used by the ML model. Of 190 predicted superconductors in a given prediction round, only 39 compounds were calculated to be stable ( $E_{\text{overhull}} = 0.00$  eV/atom) but 83 were nearly stable ( $E_{\text{overhull}} < 0.05$  eV/atom). Stable materials and those with prior experimental reports were prioritized to increase the likelihood that targeted compounds could be successfully synthesized. Prioritizing these materials ensured that failures to observe superconductivity were indicative of the behavior of the targeted compound rather than a failure to synthesize that compound.

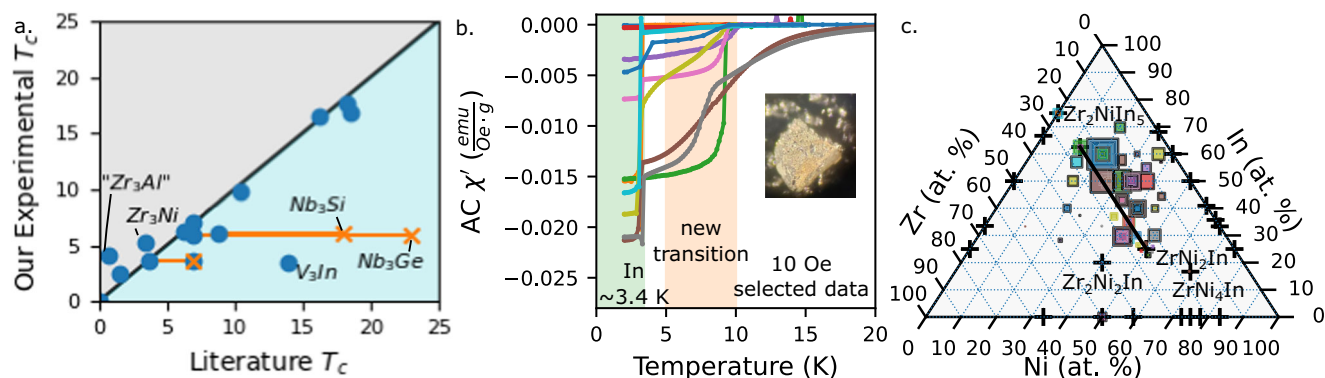
Insulating materials like  $\beta$ -ZrNiCl and the cuprates superconduct with high  $T_c$ s because they can be doped into a metallic state<sup>31</sup>. One long-running challenge for machine-learning approaches to predicting high- $T_c$  superconductivity is that large bandgap insulators incapable of superconductivity tend to be given overweighted classification scores, likely due to the high  $T_c$ s of the cuprates<sup>16</sup>. Therefore, metals and easily doped materials were favored for testing. Similarly, for some predicted metals, we investigated nearby compounds with similar structures that were known in literature but were not found in MP or OQMD (e.g., Zr<sub>3</sub>Fe<sub>4</sub>Sn<sub>4</sub> and Hf<sub>3</sub>Fe<sub>4</sub>Sn<sub>4</sub><sup>32,33</sup>) and isostructural compounds with promising band structures (e.g., ZrNi<sub>2</sub>In).

Since the  $T_c$ s of compounds are very sensitive to alloy disorder and lattice parameter, we explored several compositions near each prediction<sup>34</sup>. We also considered the ease and safety of synthesizing the target materials (e.g., by excluding extremely high-pressure syntheses). Powder X-ray diffraction (XRD) was used to ensure that the target material was successfully synthesized and temperature-dependent AC magnetic susceptibility was used to screen for superconductivity. Superconductors are perfectly diamagnetic below their  $T_c$  with minimal applied field.

### Material candidate experimental verification

To illustrate the sensitivity of experimentally-measured  $T_c$ s to processing conditions, we made and tested samples with A<sub>3</sub>B stoichiometry (Fig. 3a), including many known superconductors from the A15 family<sup>35</sup>. Similar compositional sensitivity is common in other systems beyond A15 compounds. For example, as  $x$  varies between 0 and 0.35, La<sub>2- $x$</sub> Sr <sub>$x$</sub> CuO<sub>4</sub> can vary from not superconducting to having a  $T_c$  up to 36 K<sup>15</sup>. Our experiments show that high-throughput synthesis and characterization techniques can reliably and quickly screen systems for superconductivity.





**Fig. 3 Experimental data for feedback and discovery.** **a** Evaluation of our high-throughput synthesis of compounds with  $A_3B$  stoichiometry (including  $A15$  compounds) demonstrates the effects of processing on the measured  $T_c$  and our ability to positively identify superconductors quickly. For the superconductor in the Zr-In-Ni phase diagram, samples of various compositions were tested (**b** and **c**). The size of the datapoints in (**c**) is the fraction of the superconducting phase present as estimated from the magnitude of the transition in magnetization between 5 and 10 K (orange region). This transition was distinct from the indium-related transition (green). The compositions of samples with the strongest superconducting signals cluster near the composition “ $ZrNiIn_4$ ”. The metastability of the superconducting phase precluded isolation as a single phase.

Optimization of many superconducting phases requires much lower-throughput techniques for preparing phase-pure and fully-superconducting samples.

Using this closed-loop method and high-throughput synthesis, we re-discovered five known superconductors that were not represented in the ML training dataset. A list of these is found in Table 1. Alongside these successful predictions, the ML model also returned compositions that experts could readily identify as not superconducting candidates. Therefore, it was important to compare the successful prediction rates of the combined human expert-machine approach and the machine-only approach. If one considers all predictions (including those not identified as promising by the human in the loop), the rate of discovery is 5/190 (2.6%), comparable to expert-driven success rates of (3%)<sup>10</sup>. When materials that experts quickly identified as not realistic superconductors were excluded (the human-machine combined approach), the successful prediction rate rose to 5/65 (7.5%), more than double that of previous expert-driven approaches<sup>10</sup>. This is particularly remarkable given the chemical diversity in the predicted candidates.

We were then able to use this ML model to discover unreported superconductors. Specifically, we find a superconducting phase in the Zr-In-Ni system, with a  $T_c$  of ~9 K (Fig. 3b, c and Extended Data) and approximate composition  $ZrNiIn_4$ . No other known elements, binaries or ternaries in the Zr-In-Ni system would explain a superconducting transition temperature this high and the elements and binaries have been extensively investigated<sup>12,35–37</sup>. Unfortunately, the phase responsible for superconductivity is extremely metastable, and we have not yet found a synthesis route to obtain it in single phase form (see SI).

## CONCLUSIONS

We have presented the first ever “closed-loop” ML-based directed discovery of a superconductor with experimental verification (within the Zr-Ni-In system), identified two additional systems of interest (Zr-Cu-In and Zr-Fe-Sn), and rediscovered five others not represented in our ML training set.

Past revolutionary discoveries tended to happen by serendipity, finding something in material families outside of what was known at the time. Our approach, relying only on stoichiometry and a measure of “distance” from what is currently known, is more likely to find unreported materials of interest and a sense of where unexplored but promising materials lie compared to ML-guided approaches that proceed within only a given family of materials.

This approach improves performance with experience, in that with every closing of the loop, the ML model undergoes feedback and refinement, enabling efficient exploration of materials space. These improvements ultimately will reduce the cost of materials development and discovery. The success of this approach has been demonstrated by discoveries and rediscoveries coming from vastly different families, illustrating the potential of this tool for the discovery of materials with targeted properties. This methodology can be expanded to target more than one desired property, and applied to domains beyond superconductors as long as a mechanism for new data acquisition based on ML-based predictions can be leveraged.

Further, we engaged in only a small number of total prediction/experimental measurement iterations; to maximize the superconducting transition temperatures of superconductors discovered over further iterations, we can use acquisition functions developed for Bayesian Optimization<sup>38,39</sup>. Our approach retains a human-in-the-loop for synthesizing and characterizing materials, but further automation is possible, involving, e.g., ML systems selecting experiments to be conducted, or robot-powered self-driving laboratories<sup>40–42</sup>. Thus we demonstrate a viable approach of these methods to accelerate materials discovery.

## METHODS

### Data

Our initial data source containing the superconducting transition temperature,  $T_c$  of many known compounds is the SuperCon database<sup>20</sup>, published by the Japanese National Institute for Materials Science. More details and analyses about SuperCon are available in the SI.

In this work, we use the version of SuperCon released by Stanev et al.<sup>12</sup>, available online. This contains 16,414 material compositions and associated critical temperature measurements. However, some of these compositions are invalid (e.g.,  $Y_2C_2Br_0.511.5$ ) and were removed prior to analysis. Our final training dataset has 16,304 valid compositions. In the Extended Data and the SI, we give additional detail about our training dataset. Supplementary Fig. 1 shows the distribution of  $T_c$  values in our training data—note that the distribution is weighted toward low- $T_c$  compositions.

We use MP<sup>21</sup> and OQMD<sup>22</sup> as the set of candidates to screen with ML for superconducting potential. MP and OQMD are some of the largest public sets of computational materials data. Their records contain full crystallographic information for material structures, along with some associated electronic and mechanical properties (but not, importantly,  $T_c$ ). We scraped MP for material

records present in it as of October 2020 using the `MPRester` class from the `pymatgen`<sup>43</sup> package, obtaining 89,341 unique compositions. We later downloaded the entire OQMD v1.4 database, obtaining 252,978 unique compositions. The Extended Data contains a table of MP and OQMD material IDs used in this study.

### Computational methods and uncertainty

RooSt<sup>27</sup> is a graph neural network<sup>44</sup> that relates material composition to properties by applying a message-passing scheme<sup>45</sup> to a weighted graph representation of the composition's stoichiometry, producing a real-valued embedding vector. To make a prediction, this embedding is then passed through a feedforward network.

In this work, we make use of the publicly-available implementation of RooSt, which is implemented in PyTorch<sup>46</sup>. Furthermore, we use the default hyperparameters recommended by the RooSt authors, including basing the initial species representation vectors on the *matscholar* embedding<sup>47</sup>. Since we seek materials likely to be high- $T_c$  superconductors, and we expect RooSt's classification model to poorly generalize on out of distribution data, we filter for materials predicted to be in the highest  $T_c$  tertile ( $T_c \geq 20$  K) with a classification score of at least 0.66 (see SI).

RooSt models incorporate two sources of uncertainty in their  $T_c$  predictions: We account for aleatoric uncertainty (randomness of input data) by letting a model estimate a mean and standard deviation for each label's logit<sup>48</sup>, and we incorporate epistemic uncertainty (error in the model's result, itself) by averaging over an ensemble of independently trained RooSt models<sup>49</sup>.

### Problem formulation

We formulate our prediction problem as an uncertainty-aware classification task. As shown in Supplementary Information, the distribution of  $T_c$  values in SuperCon is skewed, with a large number of materials having  $T_c$ s close to 0 K. Although we could have used a regression approach and had models estimate  $T_c$  directly, the skewed and heavy-tailed  $T_c$  distribution instead prompted us to discretize  $T_c$  into three categories, based roughly on tertiles: materials with a measured  $T_c$  less than 2 K, materials with a  $T_c$  between 2 K and 20 K, and materials with a  $T_c$  above 20 K. This is similar to earlier work by Stanev et al.<sup>12</sup>, who use a two-stage prediction approach where they first classify whether a material has a  $T_c$  of greater than 10 K. Depending on the specifics of the target property, our closed-loop discovery process can be used with other ML prediction formulations as well.

In this work, we characterize the similarity between material compositions using both the RooSt latent embedding (for predicting material properties) and via Euclidean distance applied to a material composition's Magpie<sup>30</sup> representation, for determining if superconductor candidates are not sufficiently different than known superconductors to be considered a discovery. The choice of metric is not critical, as it is imposed simply to help broaden the range of materials space explored. Other works have considered alternative mechanisms for material similarity, just as using representations based on element fractions<sup>50</sup> or the earth mover's distance<sup>51</sup>. Our discovery process does not rely on use of a specific similarity measure and can adopt other measures as desired.

### Model validation

SuperCon provides data as a validation experiment for our model—can RooSt successfully predict the  $T_c$  tertile of unknown materials? We evaluate this question in two settings; the first under a standard uniform cross-validation (Uniform-CV) split of SuperCon, and the second with the LOCO-CV strategy<sup>26</sup>. In this approach, we apply  $K$ -means clustering to the Magpie<sup>30</sup> representation of SuperCon and then train  $K$  RooSt models, iteratively

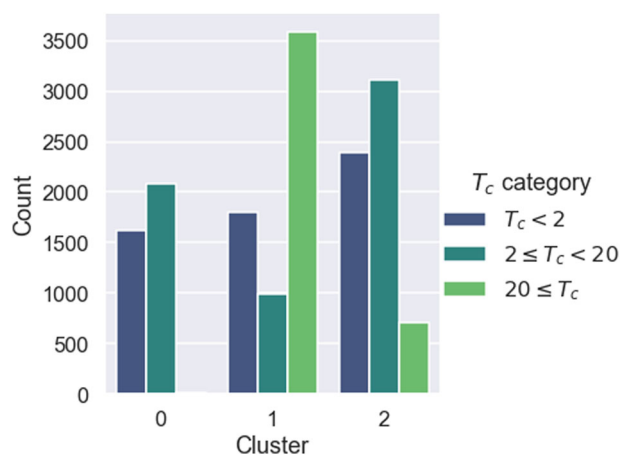
holding out each cluster as a test set. Since the clustering will put materials that are similar to each other in the same cluster, LOCO-CV is a better proxy for assessing how well our model will perform when used to identify superconductor candidates in MP.

In this study, we set  $K=3$  for the clustering and summarize cluster characteristics in Table 3 and Fig. 4. Note that even this simple clustering procedure has produced inter-cluster heterogeneity—e.g., Cluster 0 is significantly smaller than the other clusters, and Cluster 1 has the bulk of the  $20 \leq T_c$  superconductors.

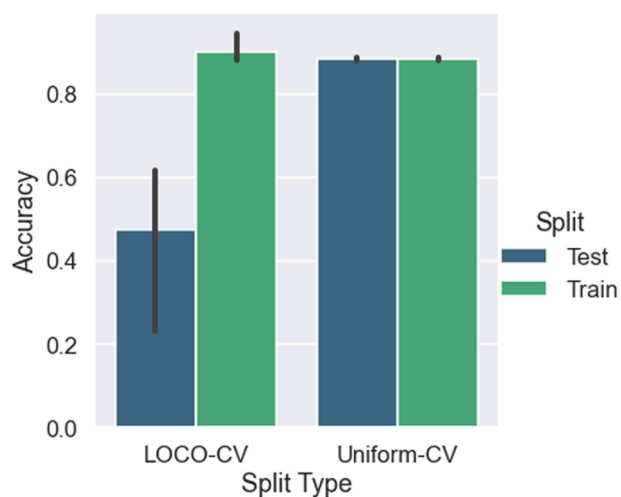
**Table 3.** LOCO-CV clustering.

Cluster	Size
0	3723
1	6377
2	6204

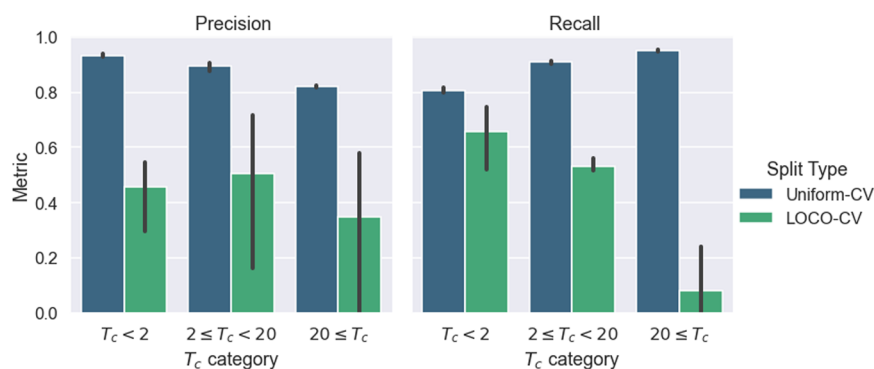
Sizes of clusters used in the LOCO-CV study, obtained from Stanev et al.<sup>12</sup>'s version of SuperCon.



**Fig. 4** Statistics of  $T_c$ . Statistics of  $T_c$  across clusters used in the LOCO-CV study, obtained from Stanev et al.<sup>12</sup>'s version of SuperCon.



**Fig. 5** Training vs. test accuracy. Training and test set accuracies for uniform cross-validation (Uniform-CV) vs. LOCO-CV, averaged over each fold and cluster. Bars show 95% confidence intervals for the standard error of the mean estimate. The model severely overfits in the LOCO-CV case, and its test set accuracy is much more cluster-dependent and variable.



**Fig. 6 Test set precision and recall.** Test set precision and recall analysis for each  $T_c$  category for the uniform vs. LOCO-CV study, averaged over each fold and cluster. Bars show 95% confidence intervals for the standard error of the mean estimate. The model's metrics are much more variable and cluster-dependent for the LOCO-CV model.

**Table 4.** A summary of the closed-loop iterations.

Loop iteration	Training Data	Candidate Data
1	SuperCon	Materials Project
2	SuperCon and measurements from iteration 1	Materials Project
3	SuperCon and measurements from iterations 1 and 2	Materials Project and OQMD
4	SuperCon and measurements from iterations 1, 2, and 3	Materials Project and OQMD

See Table 1 for discoveries made in each loop iteration, and see "Data" for the size of each dataset.

In Figs. 5 and 6, we show the results of our study. In the Uniform-CV setting, our model does well—it shows little evidence of overfitting and performs well for all three  $T_c$  categories. However, in LOCO-CV, performance degrades significantly and is also much more variable, based on what cluster is being used as the test set. Our result here echoes<sup>12</sup>, who show that models trained only on iron-based superconductors fail to accurately predict properties of cuprates, and vice versa.

These results indicate that we should not expect an ML model trained only on SuperCon to consistently identify superconductors in out-of-distribution data, and, as points in SuperCon are more similar to each other than points in MP and OQMD (Fig. 2), the LOCO-CV results here are optimistic compared to our actual problem of interest. This motivates our need for multiple iterations of model training, candidate selection, candidate synthesis, and model retraining.

### Closed-loop discovery process

The initial loop iteration used Stanev et al.'s version of SuperCon<sup>12</sup> as training data ("Data"). After the  $T_c$ -prediction model was trained, candidates were selected from MP<sup>21</sup> based on predicted scores ("Computational Methods and Uncertainty"). The second loop iteration used SuperCon, as well as additional measurements from the first loop, as training data, and it again used MP as the set of possible candidates. The third and fourth loops again used prior iterations' measurements as supplementary training data, but they also combined OQMD<sup>22</sup> to MP to obtain the set of possible candidates. The number of materials synthesized and characterized per loop iteration varied across loops, based on domain expert intuition and feasibility of synthesis. This process is summarized in Table 4.

### Experiment

To synthesize compounds in a medium-throughput manner, arc melting and solid state techniques were used. The standard sample size was 500–700 mg. A list of precursors used in this

project is found in Supplementary Table 1 in the SI and details of the synthetic procedures are found in the SI. Additional heat treatments were performed on an as-needed basis when isolating superconducting phases.

Powder XRD patterns were collected at room temperature on the as-melted samples using a Bruker D8 Focus powder diffractometer with Cu-K $\alpha$  radiation ( $\lambda_{K\alpha,1} = 1.540596$  Å,  $\lambda_{K\alpha,2} = 1.544493$  Å), Soller slits, and a LynxEye detector to verify the presence of the target phase. We measured from  $2\theta = 5^\circ$ – $60^\circ$  with a step size of  $0.018563^\circ$  over 4 min as an initial screen. When gathering XRD patterns of samples in preparation for Rietveld refinement, 4 h measurements were performed from  $2\theta = 5^\circ$ – $120^\circ$  with a step size of  $0.01715^\circ$ .

AC-susceptibility measurements were conducted using either a Quantum Design Magnetic Properties Measurement (MPMS) System ( $H_{DC} = 10$  Oe,  $H_{AC} = 1 - 3$  Oe, 900 Hz) or a Quantum Design Physical Properties Measurement System ( $H_{DC} = 10$  Oe,  $H_{AC} = 3$  Oe, 1 kHz), measuring  $T \geq 2$  K. Since prior density function theory (DFT) calculations<sup>52</sup> suggested that CaAg<sub>2</sub>Ge<sub>2</sub> would superconduct near  $T = 1.5$  K, we used the <sup>3</sup>He option with the MPMS to measure from 0.4 K to 1.7 K for that sample in addition to our standard measurement above 2 K.

### DATA AVAILABILITY

The data used to train the models used in the work were acquired from publicly available sources, and those acquired experimentally according to the procedures explained in the paper. The values acquired experimentally as part of this work are available at <https://doi.org/10.34863/w3xm-w506>.

Received: 25 April 2023; Accepted: 14 September 2023;  
Published online: 05 October 2023

### REFERENCES

1. National Research Council, *Frontiers in Crystalline Matter*. Washington, DC: The National Academies Press <https://doi.org/10.17226/12640> (2009).

2. de Pablo, J. J. et al. New frontiers for the materials genome initiative. *npj Comput. Mater.* **5**, 41 (2019).
3. Baird, S. G., Diep, T. Q. & Sparks, T. D. Discover: a materials discovery screening tool for high performance, unique chemical compositions. *Digit. Discov.* **1**, 226–240 (2022).
4. National Science and Technology Council, Materials Genome Initiative Strategic Plan: a Report by the Subcommittee on the Materials Genome Initiative Committee on Technology of the National Science and Technology Council. <https://www.mgi.gov/sites/default/files/documents/MGI-2021-Strategic-Plan.pdf> (2021).
5. Mandrus, D. Gifts from the superconducting curiosity shop. *Front. Phys.* **6**, 347–349 (2011).
6. Zhao, H. et al. Cascade of correlated electron states in the kagome superconductor  $\text{CsV}_3\text{Sb}_5$ . *Nature* **599**, 216–221 (2021).
7. Li, Y., Xu, X., Lee, M.-H., Chu, M.-W. & Chien, C. L. Observation of half-quantum flux in the unconventional superconductor beta- $\text{Bi}_2\text{Pd}$ . *Science* **366**, 238–241 (2019).
8. Mather, J. C. Super photon counters. *Nature* **401**, 654–655 (1999).
9. Grant, P. Rehearsals for prime time. *Nature* **411**, 532–533 (2001).
10. Hosono, H. et al. Exploration of new superconductors and functional materials, and fabrication of superconducting tapes and wires of iron pnictides. *Sci. Technol. Adv. Mater.* **16**(3) <https://doi.org/10.1088/1468-6996/16/3/033503> (2015).
11. Hirsch, J. Correlations between normal-state properties and superconductivity. *Phys. Rev. B* **55**, 9007–9024 (1997).
12. Stanev, V. et al. Machine learning modeling of superconducting critical temperature. *npj Comput. Mater.* **4**, 29 (2018).
13. Zeng, S. et al. Atom table convolutional neural networks for an accurate prediction of compounds properties. *npj Comput. Mater.* **5**, 84 (2019).
14. Konno, T. et al. Deep learning model for finding new superconductors. *Phys. Rev. B* **103**, 014509 (2021).
15. Roter, B. & Dordevic, S. V. Predicting new superconductors and their critical temperatures using machine learning. *Phys. C: Supercond.* **575**, 1353689 (2020).
16. Quinn, M. R. & McQueen, T. M. Identifying new classes of high temperature superconductors with convolutional neural networks. *Front. Electron. Mater.* **2**, 1–12 (2022).
17. Hoffmann, N., Cerqueira, T. F. T., Schmidt, J. & Marques, M. A. L. Superconductivity in antiperovskites. *npj Comput. Mater.* **8**, 150 (2022).
18. Seegmiller, C.C., Baird, S.G., Sayeed, H.M., Sparks, T.D. Discovering chemically novel, high-temperature superconductors <https://doi.org/10.26434/chemrxiv-2023-8t8kt-v3> (2023).
19. Lookman, T., Balachandran, P. V., Xue, D. & Yuan, R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput. Mater.* **5**, 21 (2019).
20. SuperCon (2008). <https://supercon.nims.go.jp/> Accessed 2021
21. Jain, A. et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
22. Saal, J. E., Kirklin, S., Aykol, M., Meredig, B. & Wolverton, C. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). *JOM* **65**, 1501–1509 (2013).
23. Gulrajani, I., Lopez-Paz, D. In search of lost domain generalization. In: International Conference on Learning Representations <https://openreview.net/forum?id=IQdXeXDoWtl> (2021).
24. Liu, R. & Wallqvist, A. Molecular similarity-based domain applicability metric efficiently identifies out-of-domain compounds. *J. Chem. Inf. Model.* **59**, 181–189 (2019).
25. Bemis, G. W. & Murcko, M. A. The properties of known drugs. 1. molecular frameworks. *J. Med. Chem.* **39**, 2887–2893 (1996).
26. Meredig, B. et al. Can machine learning identify the next high-temperature superconductor? examining extrapolation performance for materials discovery. *Mol. Syst. Des. Eng.* **3**, 819–825 (2018).
27. Goodall, R. E. A. & Lee, A. A. Predicting materials properties without crystal structure: deep representation learning from stoichiometry. *Nat. Comm.* **11**, 6280 (2020).
28. Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
29. Park, C. W. & Wolverton, C. Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. *Phys. Rev. Mater.* **4**, 063801 (2020).
30. 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).
31. Fournier, P. T' and infinite-layer electron-doped cuprates. *Phys. C: Supercond.* **514**, 314–338 (2015).
32. Calta, N. P. & Kanatzidis, M. G.  $\text{Hf}_3\text{Fe}_4\text{Sn}_4$  and  $\text{Hf}_3\text{Fe}_{4-x}\text{Sn}_{10+x}$ : two stannide intermetallics with low-dimensional iron sublattices. *J. Solid State Chem.* **236**, 130–137 (2016).
33. Savidan, J. C., Joubert, J. M. & Toffolon-Masclet, C. An experimental study of the Fe-Sn-Zr ternary system at 900°C. *Intermetallics* **18**, 2224–2228 (2010).
34. Matthias, B. T., Geballe, T. H., Willens, R. H., Corenzwit, E. & Hull Jr, G. W. Superconductivity of  $\text{Nb}_3\text{Ge}$ . *Phys. Rev.* **139**, 1501–1503 (1965).
35. Matthias, B. T., Geballe, T. H. & Compton, V. B. Superconductivity. *Superconductivity* **35**, 1–22 (1963).
36. Shaw, R. W., Mapother, D. E. & Hopkins, D. C. Critical fields of superconducting tin, indium, and tantalum. *Phys. Rev.* **120**, 88–91 (1960).
37. Berger, L.I., Roberts, B.W. Properties of Superconductors. In: CRC Handbook of Chemistry and Physics, 102nd edn, (eds. Rumble, J) Ch. 12 (CRC Press (Taylor and Francis), Boca Raton, FL) (2021).
38. Attia, P. M. et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **578**, 397–402 (2020).
39. Zhang, Y., Apley, D. W. & Chen, W. Bayesian optimization for materials design with mixed quantitative and qualitative variables. *Sci. Rep.* **10**, 4924 (2020).
40. Coley, C. W., Eyke, N. S. & Jensen, K. F. Autonomous discovery in the chemical sciences part 1: progress. *Angew. Chem. Int. Ed.* **59**, 22858–22893 (2020).
41. Stach, E. et al. Autonomous experimentation systems for materials development: a community perspective. *Matter* **4**, 2702–2726 (2021).
42. MacLeod, B. P. et al. Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* **6**, 8867 (2020).
43. 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).
44. Wu, Z. et al. A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **32**, 4–24 (2021).
45. Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., Dahl, G.E. Neural message passing for quantum chemistry. In: *Proc. 34th International Conference on Machine Learning - Volume 70. ICML'17*, pp. 1263–1272. JMLR.org, Sydney, Australia <https://doi.org/10.5555/3305381.3305512> (2017).
46. Paszke, A. et al. Pytorch: An imperative style, high-performance deep learning library. In: (eds. Wallach, H.M., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E.B., Garnett, R.) *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, December 8–14, 2019, Vancouver, BC, Canada*, pp. 8024–8035 <https://proceedings.neurips.cc/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf> (2019).
47. Tshitoyan, V. et al. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* **571**, 95–98 (2019).
48. Nix, D.A., Weigend, A.S. Estimating the mean and variance of the target probability distribution. In: *Proc. 1994 IEEE International Conference on Neural Networks (ICNN'94)*, vol. 1, pp. 55–601 <https://doi.org/10.1109/ICNN.1994.374138> (1994).
49. Lakshminarayanan, B., Pritzel, A., Blundell, C. Simple and scalable predictive uncertainty estimation using deep ensembles. In: *Proc. 31st International Conference on Neural Information Processing Systems. NIPS'17*, pp. 6405–6416. Curran Associates Inc., Red Hook, NY, USA (2017).
50. Jha, D. et al. Elemnet: deep learning the chemistry of materials from only elemental composition. *Sci. Rep.* **8**, 17593 (2018).
51. Hargreaves, C. J., Dyer, M. S., Gaultois, M. W., Kurlin, V. A. & Rosseinsky, M. J. The earth mover's distance as a metric for the space of inorganic compositions. *Chem. Mater.* **32**, 10610–10620 (2020).
52. Sinaga, G.S., Utimula, K., Nakano, K., Hongo, K., Maezono, R. First principles calculations of superconducting critical temperature of  $\text{ThCr}_2\text{Si}_2$ -Type Structure. *arXiv* <https://doi.org/10.48550/ARXIV.1911.10716> (2019).
53. Todorov, I. et al. Topotactic redox chemistry of NaFeAs in water and air and superconducting behavior with stoichiometry change. *Chem. Mater.* **22**, 3916–3925 (2010).
54. Hagino, T. et al. Superconductivity in spinel-type compounds  $\text{CuRh}_2\text{S}_4$  and  $\text{CuRh}_2\text{Se}_4$ . *Phys. Rev. B* **51**, 12673–12684 (1995).
55. Hiramatsu, H. et al. Water-induced superconductivity in  $\text{SrFe}_2\text{As}_2$ . *Phys. Rev. B* **80**, 2–5 (2009).
56. Pamuk, B., Mauri, F. & Calandra, M. High- $T_c$  superconductivity in weakly electron-doped  $\text{HfNiCl}$ . *Phys. Rev. B* **96**, 1–7 (2017).
57. Si, J. et al. Unconventional superconductivity induced by suppressing an Iron-Selenium-Based Mott Insulator  $\text{CsFe}_{4-x}\text{Se}_4$ . *Phys. Rev. X* **10**, 41008 (2020).
58. Ying, J., Lei, H., Petrovic, C., Xiao, Y. & Struzhkin, V. V. Interplay of magnetism and superconductivity in the compressed Fe-ladder compound  $\text{BaFe}_2\text{Se}_3$ . *Phys. Rev. B* **95**, 1–5 (2017).
59. Yamauchi, T., Hirata, Y., Ueda, Y. & Ohgushi, K. Pressure-Induced Mott Transition Followed by a 24-K Superconducting Phase in  $\text{BaFe}_2\text{S}_3$ . *Phys. Rev. Lett.* **115**, 1–5 (2015).
60. McInnes, L., Healy, J., Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction <https://arxiv.org/abs/1802.03426> (2020).

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge internal financial support from the Johns Hopkins University Applied Physics Laboratory's Independent Research & Development (IR&D) Program for funding this work. The MPMS3 system used for magnetic



characterization was funded by the National Science Foundation, Division of Materials Research, Major Research Instrumentation Program, under Award #1828490. T.M.M. acknowledges support of the David and Lucile Packard Foundation.

## AUTHOR CONTRIBUTIONS

C.D.S., I.M., C.D.P., E.G., K.M., and T.M.M. contributed to the conception of the work. E.A.P., B.W., T.M.M., E.G., and I.M. contributed to the setup and design of experiments. E.A.P., E.H., I.M., and B.W. synthesized samples. A.N. created the ML model with some contributions from M.J.P., C.D.P., and C.R.R. J.D., N.Q.L., C.C., and C.D.P. contributed to preparation of the datasets and integrating them with the workflow. E.A.P., A.N., C.D.S., K.M. and T.M.M. heavily contributed to the writing and revising of the manuscript. E.H., E.A.P., E.G., C.C., T.M.M. and B.W. collected and analyzed experimental data. E.A.P., C.D.P., A.N., T.M.M., and G.B. introduced new ways of visualizing the data. C.D.S., C.R.R., and A.L. provided guidance on research direction and communication of results. C.R.R. advised the team on appropriate ML techniques for property prediction and A.L. advised the team on material structure-property relationships. C.D.S. was involved in all aspects of this project, providing overall leadership and helping to troubleshoot technical challenges.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-023-01131-3>.

**Correspondence** and requests for materials should be addressed to Elizabeth A. Pogue, Tyrel M. McQueen or Christopher D. Stiles.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023