Prediction of large magnetic moment materials with graph neural networks and random forests

Sékou-Oumar Kaba [1,2,*] Benjamin Groleau-Paré,[3] Marc-Antoine Gauthier,[3]
A.-M. S. Tremblay [3] Simon Verret,[1,3] and Chloé Gauvin-Ndiaye [3]

[1]*Mila–Quebec Artificial Intelligence Institute & IVADO–Institut de Valorisation des Données, Montréal, Québec, Canada H2S 3H1*
[2]*School of Computer Science, McGill Univeristy, Montréal, Québec, Canada H3A 0E9*
[3]*Département de physique & Institut Quantique, Université de Sherbrooke, Québec, Canada J1K 2R1*

Magnetic materials are crucial components of many technologies that could drive the ecological transition, including electric motors, wind turbine generators, and magnetic refrigeration systems. Discovering materials with large magnetic moments is therefore an increasing priority. Here, using state-of-the-art machine learning methods, we scan the Inorganic Crystal Structure Database (ICSD) of hundreds of thousands of existing materials to find those that are ferromagnetic and have large magnetic moments. Crystal graph convolutional neural networks (CGCNNs), materials graph network (MEGNet), and random forests are trained on the Materials Project database that contains the results of high-throughput density functional theory (DFT) predictions. For random forests, we use a stochastic method to select nearly 100 relevant descriptors based on chemical composition and crystal structure. This gives results that are comparable to those of neural networks. Our findings suggests that magnetic properties are intrinsically more difficult to predict than other DFT-calculated properties. The comparison between the different machine learning approaches gives an estimate of the errors for our predictions on the ICSD database. Validating our final predictions by comparisons with available experimental data, we found 15 materials that are likely to have large magnetic moments and have not yet been studied experimentally.

I. INTRODUCTION

In recent years, materials informatics and machine learning methods have been introduced in the search for materials with specific properties, such as high-temperature superconductors [1], photovoltaics [2], radiation detector materials [3], and metallic glasses [4]. These methods have the advantage of allowing to explore sets of materials that would be prohibitively large for conventional theoretical methods or experiments. Though traditional machine learning methods such as tree-based algorithms [5,6], kernel methods [6,7], support vector machines [6], and multilayer perceptrons [6] have shown some success in the prediction of magnetic properties, the frequent introduction of new large materials databases [8–11] has also enabled the development of sophisticated neural networks for this type of application. In particular, recent Graph Neural Networks algorithms (GNNs) have recently been shown to obtain state-of-the-art performance on benchmark tasks [12–14]. These methods have been shown to successfully predict formation energy, band gap, and bulk modulus with error magnitudes similar to those of density functional theory (DFT) calculations, but have not yet been used for the study of magnetic properties.

Demand for strong permanent magnets for technological applications is rising [15]. This is closely related to the fact that many countries are looking to transition away from fossil fuels to more sustainable energy sources. Indeed, one of the main drivers of demand for permanent magnets is the production of motors for hybrid and electric vehicles, which are rapidly gaining popularity. Another growing application is wind turbine generators. For most applications, $Nd_2Fe_{14}B$ is the material of choice. However, discovering rare-earth free permanent magnets would be highly desirable for environmental and economic reasons.

Materials that have a large magnetic moment per mass unit, but that are not permanent magnets, still have multiple applications of interest. A promising application for such materials is magnetic refrigeration. Magnetic refrigeration is a technology based on the magnetocaloric effect, through which the temperature of a magnetic material varies with the adiabatic application of a magnetic field [16–18]. Because it requires the use of solid state materials instead of gaseous refrigerants, magnetic refrigeration is a more environmentally friendly technology than traditional refrigeration. However, it requires ferromagnetic materials with a Curie temperature around room temperature that also have many specific properties, such as a low specific heat and a high electrical resistivity. The reference materials that exhibit a large magnetocaloric effect around room temperature are gadolinium and some Gd-based alloys such as $Gd_5Si_2Ge_2$ [19,20]. These materials are expensive, which limits their commercial use, and are metallic, which diminishes energy efficiency due to heat loss.

Often, the search for new materials that exhibit specific properties is done through trial and error, limiting the number of materials that can be studied both theoretically and
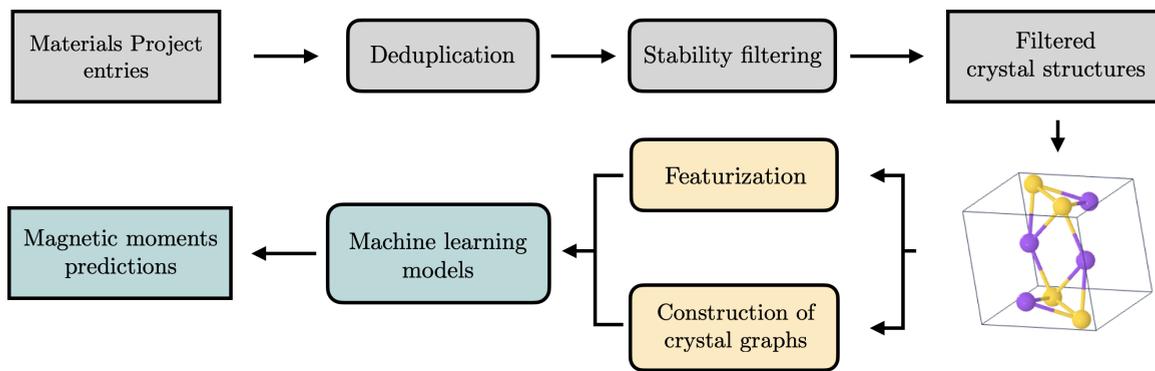
*kabaseko@mila.quebec

FIG. 1. Data flow for model training, including cleaning, preprocessing, and prediction steps.

experimentally. Considering the challenges that surround the design of new materials for magnetic refrigeration and other applications, we investigate the usefulness of machine learning for materials in the study of magnetic properties as a first step towards the discovery of new materials.

In this paper, we assess the performance of two recently proposed neural networks, crystal graph convolutional neural networks (CGCNNs) [12] and materials graph network (MEGNet) [13], and compare it to that of the random forest, a statistical machine learning method [21–23], for the prediction of the magnetization of materials. To do so, we train our models on the Materials Project database, a DFT database frequently used as a training set in the field of machine learning for material properties [8]. We characterize the Materials Project data set and describe a preprocessing scheme based on the energy above hull to reduce bias in its distribution of magnetic orders. Figure 1 shows the steps used in training the models. Our work reveals that neural networks are comparable to random forests for the prediction of magnetization. We also find that the performance of machine learning models in fitting magnetization is comparatively worse than for other properties such as formation energy, suggesting that it is an intrinsically more difficult task. We then apply our trained models on the ICSD database, which contains around 100 000 stoichiometric experimentally studied materials [24]. We discuss the suitability of the proposed materials for the specific application to magnetic refrigeration.

In the following section, we describe the data sets. Section III explains the machine-learning methods, followed by our predictions and error estimates in Sec. IV. A discussion in Sec. IV B is followed by the conclusion in Sec. V.

II. DATA SETS

We first discuss the Material's Project database that was used to train the models and then the ICSD database of materials that we use to make predictions.

A. Materials Project

The training data used in this work comes from the Materials Project data set (V2020.06) [8]. It is one of the largest data sets obtained from high-throughput DFT calculations and has become standard in machine learning based materials studies.

The Materials Project comprises stoichiometric crystalline materials, and provides their chemical composition, relaxed structures, and a number of properties such as the formation energy, the energy above hull, the band structure, and the spontaneous unit cell magnetization. To our knowledge, this last property in Materials Project has not yet been used for machine learning applications. Our first objective is to determine whether or not the magnetization is a property that can be modeled properly with machine learning algorithms.

We note that all the calculations are initialized in the ferromagnetic configuration. Antiferromagnetic configurations can be reached in the crystal relaxation stage. However, it has been shown that this method favors ferromagnetic configurations, even when low-spin antiferromagnetic or ferrimagnetic configurations could have lower energies [25]. In 2019, a new workflow was introduced as an effort to include appropriate antiferromagnetic ground states and counter the ferromagnetic bias [25]. At this time, new ground-state calculations using this workflow were performed for about 520 materials, less than 2% of the Materials Project data set.

The most difficult materials to simulate with DFT calculations are the ones in which electronic correlations are strong [26]. This is most notable in materials that contain d or f valence electrons, such as the transition metals and the rare earths. These electrons are also the ones that participate in the magnetic properties. In the Materials Project data set, all calculations on oxides containing Co, Cr, Fe, Mn, Mo, Ni, V, and W atoms are performed with the generalized gradient approximation$+U$ (GGA$+U$) scheme that aims to better represent the electronic correlations. We still note that the magnetic ground state can be significantly influenced by the effect of electron-electron interactions beyond the GGA$+U$ approximation.

We filter the entries of the Materials Project data set to create our training set. This is because the predictions of our machine learning models can only come from the identification of patterns within this data set. It is therefore crucial to identify biases and potential obstacles to generalization in this training distribution [27]. We first remove duplicates from the data set. Indeed, in the data set there is an entry for each calculation, and not each material, which results in having entries with only marginal structural differences. We use a simple heuristic to remove duplicates from the data set. Two materials are considered similar if they share the same unit cell composition and space group. This criterion
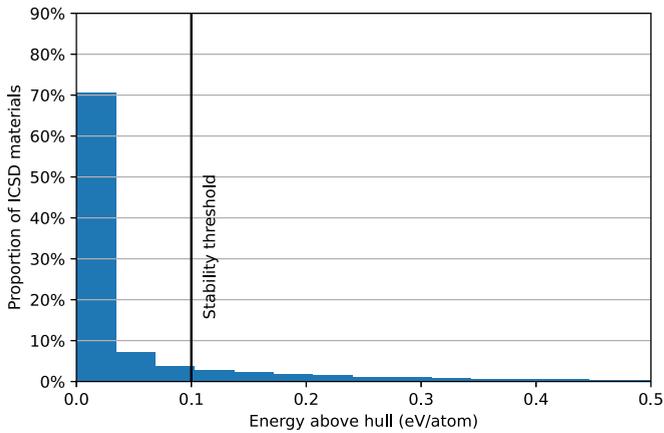
FIG. 2. Distribution of the energy above hull attribute in the Materials Project data set for the subset of entries that are also in the ICSD data set.

FIG. 4. Distribution of magnetic orders in the Materials Project data set. NM stands for nonmagnetic, FM for ferromagnetic, FiM for ferrimagnetic, and AFM for antiferromagnetic.

is conservative, but computationally tractable and allows to effectively remove duplicates.

Another important source of errors comes from the fact that DFT can relax to unstable, high-energy structures. Not only is it that these materials cannot be synthesized and will drive the distribution away from materials of interest, but they will also tend to exhibit atypical features that could confuse training. We thus choose to filter these materials out of the data set based on their energy above hull computed by DFT. The energy above hull gives the energy of decomposition of a material into a set of stable materials containing all the chemical elements of the original material and whose total formation energy is smaller. As shown in Fig. 2, the distribution of energies above hull is highly skewed towards small values for entries that also appear in ICSD (i.e., materials that have been synthesized). We choose heuristically the value of $E = 0.1$ eV/atom as a threshold for stability. This results in keeping 69% of the materials in the data set (Fig. 3). The total curated data set leaves us with 78 462 materials.

The Materials Project data set is not specifically focused on magnetic materials and includes nonmagnetic (which
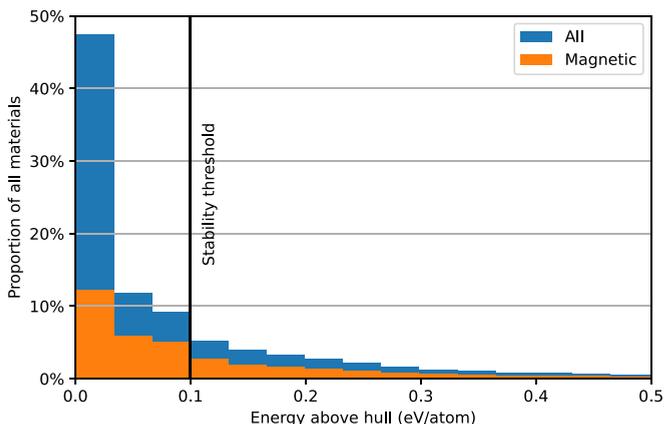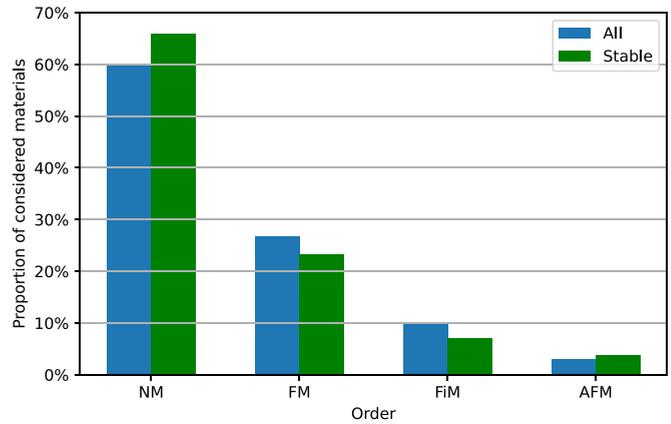
includes paramagnetic since only spontaneous magnetization is reported), ferromagnetic, and antiferromagnetic materials. For our goal of predicting magnetization, having both non-magnetic and magnetic materials is a desirable property: We want the model to identify the factors promoting strong magnetization as well as those inhibiting it. Figure 3 shows that the proportion of magnetic materials increases with the energy above hull, so eliminating large energy above hull material helps fill this criterion of having a more balanced proportion of magnetic and nonmagnetic materials.

The number of materials for each magnetic order is indicated in Fig. 4 for all materials and for those that are stable according to our criterion. We note that the procedure of filtering out unstable structures helps to alleviate the ferromagnetic bias of DFT. Figure 4 further shows that the eliminated materials are more ferromagnetic and ferrimagnetic; the proportion of antiferromagnetic materials slightly increases after filtering.

Further details on the Materials Project are provided in Appendix A of the Supplemental Material [28].

B. ICSD

Having trained models on the Materials Project, our second objective is to identify high magnetization candidates that can readily be synthesized. We use data from the Inorganic Crystal Structures Database (ICSD) to perform this task. It is currently the largest database of experimentally identified crystalline materials.

ISCD data are obtained directly from scientific publications. It includes chemical composition as well as crystal structure data for all of its entries.

Unlike the Materials Project, ICSD includes nonstoichiometric materials. Since our training distribution did not include such materials, we expect that our models will have difficulty generalizing to these materials and we have thus removed these entries from the inference data set. In addition, it is crucial to take into account that the Materials Project and ICSD data sets are not independent. The crystal structures of the entries in the Materials Project are computed starting

FIG. 3. Distribution of the energy above hull attribute in the Materials Project data set.
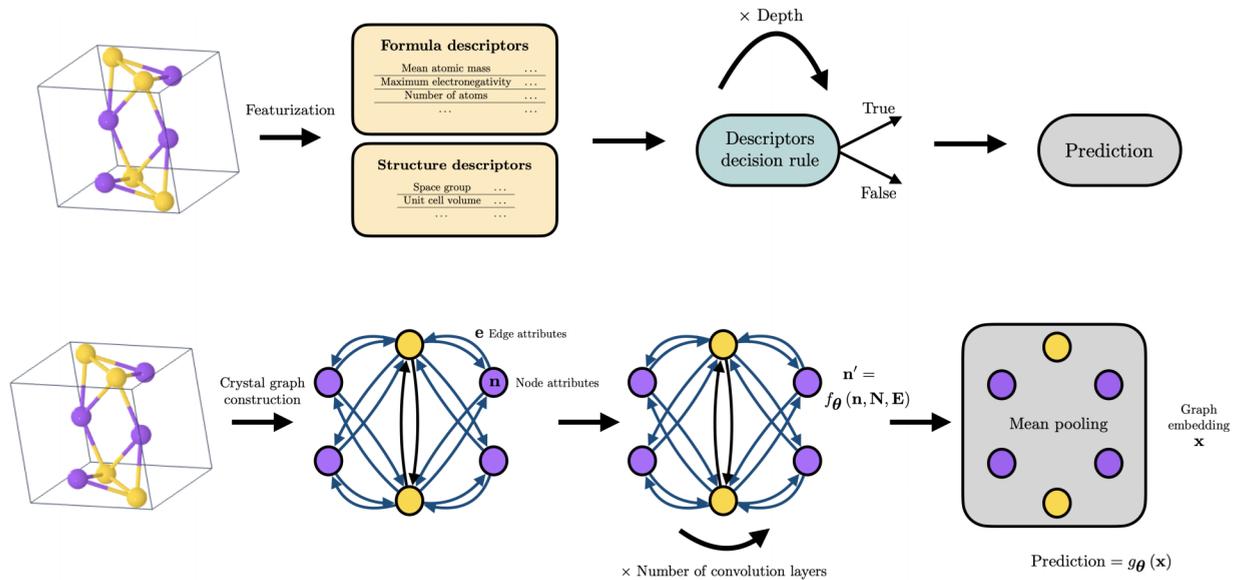
FIG. 5. Methods used for property prediction. Top panel: Illustration of a decision tree. First, descriptors are computed from the chemical formula and the structural properties. A series of decisions are then taken using the descriptors. After a number of decisions, the final prediction is made. Bottom panel: Illustration of a GNN. A graph is built from the crystal unit cell. Each atom is mapped to a node and an edge is drawn between two nodes if they share a Voronoi face. Both nodes and edges have associated embedding vectors. The graph then goes through a series of graph convolution operations (defined differently for each method) parametrized by the neural network f_θ. Finally, the features of all the nodes are averaged and the resulting vector goes through a multilayer perceptron g_θ that outputs the prediction.

from ICSD entries. Though some of the structures change significantly through the relaxation process, many of them remain sufficiently similar to the ICSD starting point to be considered as common entries between the two data sets. We have identified the common entries in both data sets and found that out of the 110 870 entries in ICSD that are stoichiometric and compatible with the methods, 23 311 materials did not have matching identifiers or formulas. Inference of magnetization was therefore performed on this subset.

III. METHODS

Once trained, machine learning algorithms have the crucial advantage of producing predictions orders of magnitude faster than simulation methods such as DFT. We can thus use them to efficiently screen candidate materials in a large database. The setup is that of a supervised learning task: Given a training set of materials with features $\mathbf{X}$ and known ground-truth targets $\mathbf{Y}$, the algorithm is tasked to learn a prediction function $\hat{\mathbf{Y}} = f_\theta(\mathbf{X})$. Training parameters θ are optimized to minimize a loss function $\mathcal{L}(\mathbf{Y}, \hat{\mathbf{Y}})$.

In the following, we describe the learning algorithms we have used. Hyperparameters are given in Appendix B of the Supplemental Material [28]. Aside from the more complex machine learning methods discussed in this paper, we also used the linear model [29]. The linear model is trained using the same descriptor space as the random forest. Since the linear model assumes a simple relation between each descriptor and the target, it is not expected to be accurate. Therefore, it establishes a performance baseline against which more elaborate methods will be compared. Figure 5 summarizes schematically the methods that we use. We explain each of them in more detail below.

A. Random forests

The random forest [30] is a tree-based machine learning algorithm that has been widely used for materials property prediction [4–6]. Random forests have yielded encouraging results for similar material design tasks, for example finding superconductors [1,31]. At the root of any tree-based method lies in the decision tree, which carries out consecutive binary splits in the descriptor space of the data (see Fig. 5). Single decision trees, however, have the major drawback of frequently overfitting the training data. Random forests go around this problem by averaging the predictions of multiple decision trees. Each tree is built differently to ensure that there is some variance in the generated forest. Randomness is implemented by choosing a subset of the descriptors to be available to the usual tree algorithm every time the algorithm makes a split. The subset is different for every split. The number of descriptors in the subset is a hyperparameter called "Available features per split." The use of random forests is motivated in our case by their relative simplicity and efficiency [32]. We use SCIKIT-LEARN's [29] implementation of the random forest methods.

The model takes descriptors handcrafted for each sample material as input. The algorithm achieves much better performance if descriptors are well adapted to the task. The properties found in both Materials Project and ICSD, such as density and crystal structure, can be employed in our case. Inspired by other works [1,31], most of the descriptors are obtained from the chemical formula of the materials: Starting from atomic properties, such as the ground-state magnetic moment, the electronegativity, the atomic mass, or the ground-state d-shell electrons, we compute the mean value, the maximum value, or the standard deviation of each of these properties to form descriptors. A number of descriptors

are also obtained from the crystal structure, for example the number of sites in the unit cell or bond lengths. We design more than 400 descriptors in this way for each material encountered. Appendix D of the Supplemental Material [28] gives details on descriptor design and provides an exhaustive list of all descriptors available to our random forest model.

Using this large descriptor space would lead to overfitting. Hence we identify a subset of descriptors that, in addition to yielding better predictions, comes with the added benefit of giving more interpretability to the model. Because of the large number of descriptors and material entries, forward and backward descriptor selection methods one would typically use for this task require an unreasonable amount of computational resources and time. We therefore design a descriptor selection scheme that mixes both forward and backward descriptor selection in order to efficiently find which descriptors are relevant for the task (details in Appendix E of the Supplemental Material [28]).

B. Graph neural networks

Graph neural networks (GNNs) are deep learning architectures that are widely used for molecular-property prediction and generation [33,34]. Their use in the context of materials is however recent. The main advantage of using GNNs for materials property prediction compared to other machine learning methods is that they only take as input a graph encoding of the material that naturally encodes structure information. The need for using handcrafted features is eliminated, as the deep model acts as a *feature extractor*. This ability to learn a representation adapted to the task at hand from the raw data has been key to the success of deep learning in a variety of domains. However, it is essential to note that this additional expressivity (the complexity and diversity of prediction functions that can be learned) comes at the cost of interpretability. It is notoriously difficult to know how a neural network selects specific features for predictions [35]. For this reason, we deem more appropriate to use neural networks in conjunction with random forests, a method that allows explicit descriptor construction and selection.

Here, we use two architectures, CGCNN [12] and MEGNet [13], that have been chosen for their performance on other properties of the Materials Project data set as well as their good training speed. For each material, the crystal unit cell is mapped to a sparse graph. This is done by associating each atom to a node and linking two nodes if they share a Voronoi face and are within a cutoff distance of 5 Å. Edges are also added if a face is shared with an atom outside the unit cell to enforce periodic boundary conditions. Around 5% of structures resulted in disconnected graphs and were discarded. When building the graphs, node features are added by taking one-hot encodings of the corresponding atomic type. This allows to capture all the information on atoms composing the unit cell. An encoding based on the distance for edges is used as well. Taking inspiration from SchNet [14], this distance is expanded on a Gaussian basis of functions, with details specified in Appendix B of the Supplemental Material [28].

Each architecture takes these labeled graphs as input and applies successive graph convolution layers to them (see Fig. 5). These convolution operations consist of each node aggregating the features of its neighbors using a learned function. The main difference between the two models revolves around the design of these layers, which is detailed in the original papers. After a number of passes in convolution layers, node and edge representations are pooled and sent to a regressor multilayer perceptron that outputs the final prediction. For each architecture, we use the original hyperparameters with a few modifications detailed in Appendix B in the Supplemental Material [28]. Training is performed using a stochastic gradient descent with the ADAM optimizer [36] as well as a learning rate scheduler.

C. Training and inference

We first train a model for each method on the Materials Project data set. Training is performed by minimizing the mean squared error (MSE) of the predicted magnetization per atom with the Materials Project target,

$$\mathcal{L}_{\mathrm{MSE}} = \frac{1}{N}\|\mathbf{Y} - \hat{\mathbf{Y}}\|_2^2, \tag{1}$$

where N is the size of the data set. We also report the mean absolute error (MAE) between predicted magnetization and ground-truth values,

$$\mathcal{L}_{\mathrm{MAE}} = \frac{1}{N}\|\mathbf{Y} - \hat{\mathbf{Y}}\|_1. \tag{2}$$

The Materials Project data set is split into a training set comprising 80% of the data, a validation set of 10%, and a test set of 10%. We compare the performance on the test of each method in Table II. We also train models to predict the formation energy given in Materials Project. Since results on this task are obtained in the original implementations of CGCNN and MEGNet, this allows us to verify that our implementations of the models perform as well as expected.

Inference is performed using the trained model on the ICSD database. As explained above, nonstoichiometric (doped) materials are eliminated since they are absent from the training set. In addition, a small fraction of the entries had incomplete crystal structure information and had to be discarded when using GNNs.

IV. RESULTS

We start by comparing the behavior of the various methods for the prediction of the magnetization. Then, we give the predictions for magnetic moments of compounds in the ICSD data set. We also discuss the importance of the random forest descriptors in Appendix F of the Supplemental Material [28].

A. Evaluation of methods

We first look at the task of predicting formation energy used as a benchmark for GNN methods. We find that our models perform on par or better than the original implementations. The difference can be attributed to the fact that we use a different version of the Materials Project data set as well as different training, validation and test set splits. Our hyperparameters are also slightly different, as detailed in Appendix B of the Supplemental Material [28]. We see

TABLE I. Performance on the test set of the various models on prediction of the formation energy per atom using the Materials Project data set.

Model	MAE (eV/atom)	MSE (eV/atom)
CGCNN (original)	0.039	
MEGNet (original)	0.028	
Linear model	0.302	0.169
Random forest	0.100	0.038
CGCNN (Ours)	0.023	0.003
MEGNet (Ours)	0.031	0.004

that neural networks outperform random forests (Table I) by a significant margin.

Then, we evaluate the different models on the magnetic moment prediction task on the Materials Project data set. Results are shown in Table II. As expected, all models outperform the baseline linear model. We find that both neural network architectures perform worse on this task than random forests. This is in strong contrast with the evaluation of the prediction accuracy of formation energy. Thus, the common assumption that deep models should perform better than models based on handcrafted descriptors does not hold for the task of predicting the magnetic moment.

The difference between magnetization and formation energy results may be understood by the fact that random forests may handle imbalance in prediction labels better than neural networks. In our case, the imbalance is caused by the bimodal distribution of magnetization values, with one mode associated with nonmagnetic and antiferromagnetic materials and the other with ferromagnetic and ferrimagnetic materials. These modes have significantly different weights in the distribution as shown in Fig. 4. It is well known that neural networks are difficult to train on imbalanced data [37]. Deep models also show a stronger tendency to overfit training data which can explain that they compare worse on MSE than MAE. This is also confirmed by the MSE on the training set in Table VII of Appendix C of the Supplemental Material [28], which is one order of magnitude smaller for neural networks than for random forests.

Finally, all methods show only a smaller improvement in performance with respect to the linear model compared to the improvement on formation energy prediction. The reasons for this difficulty of predicting magnetization with more elaborate machine learning are unclear. One could argue that this could be due to the biased distribution of magnetization data in the Materials Project. This is however not an entirely satisfying

TABLE II. Performance on the test set of the various models on prediction of the magnetic moment per atom using the Materials Project data set.

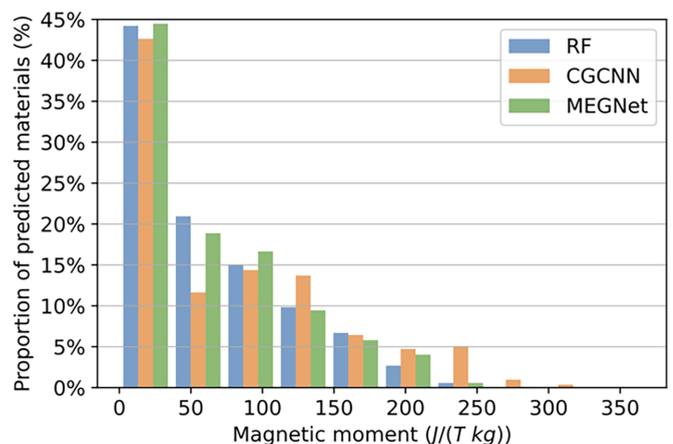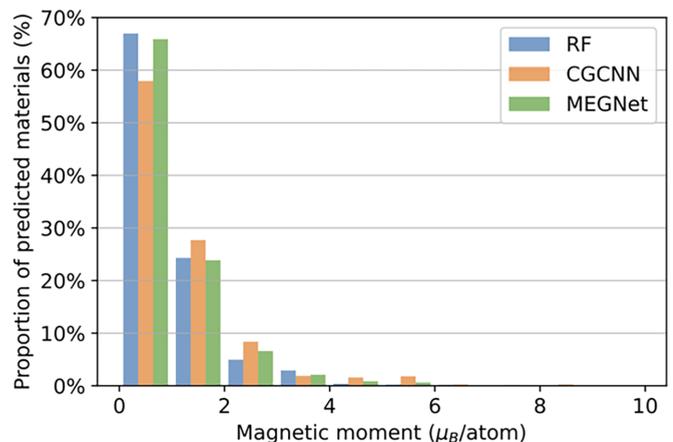| Model	MAE (μ_B/atom)	MSE (μ_B/atom)
Linear	0.100	0.030
Random Forest	0.043	0.015
CGCNN	0.052	0.026
MEGNet	0.052	0.026

FIG. 6. Distribution of the predicted magnetic moment for the three models: random forests (RF), CGCNN, and MEGNet. The top panel shows the results in units of the Bohr magneton per atom (bin size is one Bohr magneton), and the bottom panel in units of magnetization per kg.

explanation because the test distribution is just as biased as the training distribution is. Our results instead suggest that magnetization is an intrinsically more challenging property to predict than formation energy. Better predictions could then be obtained either by incorporating more physics-based biases into the models or by accumulating more data.

B. Predictions on ICSD

We now apply the trained random forest, CGCNN and MEGNet models on the ICSD data set. We first start with the disclaimer that the predictions shown below are limited in two ways: (1) by the ferromagnetic bias of the Materials Project data set and (2) by the difficulty of predicting the magnetic moment, as shown in the previous section. We use the median and standard deviation of the results from the three models to estimate the magnetization and the error. In Fig. 6, we illustrate the distribution of the predicted magnetic moments obtained from our three machine learning models. Despite the ferromagnetic bias present in the Materials Project data set, we note all three models predict a majority of materials with little to no magnetization. More precisely, CGCNN, the

TABLE III. Distribution of the magnetic orders reported experimentally for 64 materials present in our predictions on ICSD. The middle (right) column lists the results for the 32 materials with the highest (lowest) median magnetic moment per kilogram predicted by our models for which we could find experimental data.

Magnetic order reported from experiments	Number of materials (highest magnetic moments)	Number of materials (lowest magnetic moments)
FM/FiM	17 (53%) [19,38–53]	3 (9%) [54–56]
AFM	12 (38%) [51,57–67]	7 (22%) [68–74]
NM/PM	2 (6%) [75,76]	11 (34%) [77–85]
DM	0 (0%)	10 (31%) [77,86–89]
Other	1 (3%) [90]	1 (3%) [91]

random forests, and MEGNet predict that 45%, 47%, and 48% of the materials have a magnetic moment smaller than $0.5\mu_B$/atom, respectively. In Appendix G of the Supplemental Material [28], we show an extended analysis of the predictions on ICSD.

To better understand the accuracy of our predictions, we sort the results in decreasing order of the predicted magnetization per kilogram. Then, we focus on the first and last 150 materials with the highest and lowest predicted magnetic moment per mass unit from this list. We were left with a total of about 120 materials after removing duplicate entries (about 60 materials with a high magnetic moment and 60 with a zero magnetic moment). To gain an insight on the accuracy of our model, we compare the predicted magnetic properties of these materials to available experimental measurements and report our findings in Table III. Notably, we were unable to find experimental reports on the magnetic properties of about 30 of the materials that were predicted as having a high magnetic moment, indicating that our models could indeed be used for the discovery of new magnetic materials. We report our predictions for the magnetic moment of 15 of these materials in Table IV. We note that all of these materials include magnetic rare earths or magnetic transition metals. It

TABLE IV. Predictions of the largest magnetic moments per mass on the ICSD database for which no magnetic order reported from experiments could be found. Median values and standard deviations of the predictions of random forests, CGCNN, and MEGNet are shown.

Chemical formula	Moment per mass [J/(T kg)]	Rare earths
Fe_9O	240 ± 40	No
$(Mn_2O_3)_3MnSiO_3$	230 ± 50	No
$EuFe_2$	220 ± 60	Yes
$EuOF$	220 ± 20	Yes
Eu_2Cu	220 ± 40	Yes
$Mn_6O(VO_4)_3(OH)$	220 ± 80	No
EuH	220 ± 90	Yes
Fe_8C_2	210 ± 20	No
Fe_8N_2	200 ± 30	No
$GdAl$	200 ± 20	Yes
$Eu(CN_2)$	200 ± 60	Yes
$Mn_8Si_6O_{24}ClH_9$	200 ± 60	No
Gd_6Co_2Al	200 ± 50	Yes
$Mn_8O_{10}Cl_3$	200 ± 40	No
$EuBeGd_2O_5$	200 ± 80	Yes

is hence plausible that they are indeed ferromagnets with a large magnetic moment.

Coming back to Table III, we first comment on the results for the materials with high predicted magnetic moments. We find that 17 (53%) materials from this list are actually reported as ferromagnetic, or ferrimagnetic with a high magnetic moment per mass unit. The remaining materials, accounting for 47% of this sample, are found to be mostly antiferromagnetic. This analysis of a subset of our predictions highlights the challenge in predicting the magnetization from models trained of the Materials Project data set. The ferromagnetic bias in the training set may explain the discrepancy between the predicted ferromagnetic orders and the actual nonferromagnetic orders observed experimentally. This emphasizes that the discrepancies between high-throughput DFT calculations and experiments can mean that the high performance of a model trained on a DFT data set does not necessarily translate to accurate predictions when it is applied to an experimental database. We note that this conclusion applies to other predictions, and not only to the prediction of the magnetic properties because the magnetic order impacts all the ground-state properties obtained through a DFT calculations.

The comparison of our predictions for small magnetic moments to available experimental data enables a better understanding of the predictive power of our models. Indeed, for these materials, we find that only three materials (9%) from this sample are experimentally found to be ferromagnetic or ferrimagnetic, with the remaining 29 being antiferromagnetic, nonmagnetic, paramagnetic, or diamagnetic, all of which are orders with zero (or negligible) net magnetic moment. Hence, as noted before from Fig. 6, the ferromagnetic bias present in the training data set does not preclude the models from accurately predict small magnetic moments.

With regards to magnetic refrigeration, the rotating magnetocaloric effect is a promising recent research direction [92–96]. While the traditional magnetocaloric effect is due to the change in magnetization of a material under the successive application and removal of an external magnetic field, the rotating magnetocaloric effect is due to a strong magnetic anisotropy. In such anisotropic materials, the magnetization is locked in a particular crystal direction (the so-called "easy axis"). In that case, a magnetocaloric effect is seen when the material is rotated in a fixed external magnetic field. Materials that exhibit this rotating magnetocaloric effect must have a strong spin-orbit coupling, which is usually associated to elements with a large atomic number. The materials we

show in Table IV that include Eu, Gd, and In elements could potentially satisfy this requirement.

V. CONCLUSION

To identify materials with a large magnetic moment per kilogram, we have trained random forests and two state-of-the-art deep-learning graph convolutional algorithms, CGCNN and MEGNet, on the Materials Project data set. Since the discovery of large magnetic moments materials could lead to breakthroughs in magnetocaloric refrigeration, we choose this application as a use case. Magnetic properties in the training set are computed using DFT methods. The three machine-learning methods show comparable accuracy on the test sets. Differences in estimates for the mean average error and mean squared error suggest that the predicted magnetic moment per atom is accurate to better than 0.05 Bohr magneton per atom.

We used the trained models to search for candidate materials with large magnetic moments per kilogram in the ICSD database. That database contains materials that have been synthesized. Table IV lists the most promising materials. These deserve experimental attention. Our analysis also highlights some of the limitations of working with magnetic properties in the Materials Project. In particular, the ferromagnetic bias affects the screening capability of models trained on this data set. In the few cases where we could compare our predictions for ferromagnetic materials with experiment, we found that materials were indeed ferromagnetic 50% of the time. Since the rate of false negatives is low, we believe our model does not leave out many materials that are ferromagnetic. Finally, we have shown that magnetization may be intrinsically more difficult to predict than other properties. These findings provide strong arguments to motivate building a large database dedicated to the magnetic properties of materials.

Data from the Materials Project data set are freely available using their API. Documentation for this API is available online [97]. ICSD is under a proprietary license. Code for the machine learning models is available from the corresponding author upon reasonable request.

[1] Z.-L. Liu, P. Kang, Y. Zhu, L. Liu, and H. Guo, Material informatics for layered high-T_C superconductors, APL Mater. **8**, 061104 (2020).

[2] Y. Wu, J. Guo, R. Sun, and J. Min, Machine learning for accelerating the discovery of high-performance donor/acceptor pairs in non-fullerene organic solar cells, npj Comput. Mater. **6**, 120 (2020).

[3] C. Ortiz, O. Eriksson, and M. Klintenberg, Data mining and accelerated electronic structure theory as a tool in the search for new functional materials, Comput. Mater. Sci. **44**, 1042 (2009).

[4] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, A general-purpose machine learning framework for predicting properties of inorganic materials, npj Comput. Mater. **2**, 16028 (2016).

[5] G. A. Landrum and H. Genin, Application of machine-learning methods to solid-state chemistry: ferromagnetism in transition metal alloys, J. Solid State Chem. **176**, 587 (2003).

[6] T. D. Rhone, W. Chen, S. Desai, S. B. Torrisi, D. T. Larson, A. Yacoby, and E. Kaxiras, Data-driven studies of magnetic two-dimensional materials, Sci. Rep. **10**, 15795 (2020).

[7] J. J. Möller, W. Körner, G. Krugel, D. F. Urban, and C. Elsässer, Compositional optimization of hard-magnetic phases with machine-learning models, Acta Mater. **153**, 53 (2018).

[8] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, The Materials Project: A materials genome approach to accelerating materials innovation, APL Mater. **1**, 011002 (2013).

[9] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, The Open Quantum Materials Database (OQMD): Assessing the accuracy of DFT formation energies, npj Comput. Mater. **1**, 15010 (2015).

[10] C. J. Court and J. M. Cole, Magnetic and superconducting phase diagrams and transition temperatures predicted using text mining and machine learning, npj Comput. Mater. **6**, 18 (2020).

[11] S. Curtarolo, W. Setyawan, G. L. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. J. Mehl, H. T. Stokes, D. O. Demchenko, and D. Morgan, AFLOW: An automatic framework for high-throughput materials discovery, Comput. Mater. Sci. **58**, 218 (2012).

[12] T. Xie and J. C. Grossman, Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties, Phys. Rev. Lett. **120**, 145301 (2018).

[13] C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong, Graph networks as a universal machine learning framework for molecules and crystals, Chem. Mater. **31**, 3564 (2019).

[14] K. Schütt, P.-J. Kindermans, H. E. S. Felix, S. Chmiela, A. Tkatchenko, and K.-R. Müller, SchNet: A continuous-filter convolutional neural network for modeling quantum interactions, in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, edited by I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Neural Information Processing Systems Foundation, Inc. (NeurIPS), Long Beach, California, 2017), pp. 991–1001.

[15] H. Nakamura, The current and future status of rare earth permanent magnets, Scr. Mater. **154**, 273 (2018).

[16] P. Weiss and A. Piccard, Le phénomène magnétocalorique, J. Phys. Theor. Appl. **7**, 103 (1917).

[17] J. R. Gómez, R. F. Garcia, A. D. M. Catoira, and M. R. Gómez, Magnetocaloric effect: A review of the thermodynamic cycles in magnetic refrigeration, Renewable Sustainable Energy Rev. **17**, 74 (2013).

[18] A. Smith, C. R. Bahl, R. Bjørk, K. Engelbrecht, K. K. Nielsen, and N. Pryds, Materials challenges for high performance magnetocaloric refrigeration devices, Adv. Energy Mater. **2**, 1288 (2012).

[19] G. V. Brown, Magnetic heat pumping near room temperature, J. Appl. Phys. **47**, 3673 (1976).

[20] V. K. Pecharsky and K. A. Gschneidner, Jr., Giant Magnetocaloric Effect in $Gd_5(Si_2Ge_2)$, Phys. Rev. Lett. **78**, 4494 (1997).

[21] L. Breiman, Random forests, Mach. Learn. **45**, 5 (2001).

[22] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, 8th ed. (Springer, New York, 2017).

[23] A. Liaw and M. Wiener, Classification and regression by randomForest, R News **2**, 18 (2002).

[24] G. Bergerhoff, R. Hundt, R. Sievers, and I. Brown, The inorganic crystal structure data base, J. Chem. Inf. Comput. Sci. **23**, 66 (1983).

[25] M. K. Horton, J. H. Montoya, M. Liu, and K. A. Persson, High-throughput prediction of the ground-state collinear magnetic order of inorganic materials using density functional theory, npj Comput. Mater. **5**, 64 (2019).

[26] E. Pavarini, Solving the strong-correlation problem in materials, Riv. Nuovo Cimento **44**, 597 (2021).

[27] V. Gudivada, A. Apon, and J. Ding, Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations, Int. J. Adv. Softw. **10**, 1 (2017).

[28] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevMaterials.7.044407 for information on the data sets used in the study, implementation of the machine learning models, and additional results, which includes Refs. [98–102].

[29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. **12**, 2825 (2011).

[30] T. K. Ho, Random decision forests, in *Proceedings of 3rd International Conference on Document Analysis and Recognition* (IEEE, New York, 1995), Vol. 1, pp. 278–282.

[31] V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, and J. Paglione, Machine learning modeling of superconducting critical temperature, npj Comput. Mater. **4**, 29 (2018).

[32] G. Louppe, Understanding random forests: From theory to practice, Ph.D. thesis, University of Liège, 2015.

[33] D. K. Duvenaud, D. Maclaurin, J. Iparraguirre, R. Bombarell, T. Hirzel, A. Aspuru-Guzik, and R. P. Adams, Convolutional networks on graphs for learning molecular fingerprints, in *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (Neural Information Processing Systems Foundation, Inc. (NeurIPS), Montréal, Canada, 2015), pp. 2224–2232.

[34] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, Neural message passing for quantum chemistry, in *Proceedings of the 34th International Conference on Machine Learning*, edited by D. Precup and Y. W. Teh, Proceedings of Machine Learning Research (PMLR, 2017), Vol. 70, pp. 1263–1272.

[35] S. Chakraborty, R. Tomsett, R. Raghavendra, D. Harborne, M. Alzantot, F. Cerutti, M. Srivastava, A. Preece, S. Julier, R. M. Rao *et al.*, Interpretability of deep learning models: A survey of results, in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* (IEEE, New York, 2017), pp. 1–6.

[36] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.

[37] J. M. Johnson and T. M. Khoshgoftaar, Survey on deep learning with class imbalance, J. Big Data **6**, 1 (2019).

[38] X. Chen and Y. H. Zhuang, Magnetocaloric effect of $Gd_{12}Co_7$, Solid State Commun. **148**, 322 (2008).

[39] J. Crangle and G. M. Goodman, The magnetization of pure iron and nickel, Proc. R. Soc. London, Ser. A **321**, 477 (1971).

[40] H.-M. Hong, Y.-J. Kang, J. Kang, E.-C. Lee, Y.-H. Kim, and K. J. Chang, Effect of chemical bonding on the magnetic stability and magnetic moment in Mn-based binary compounds, Phys. Rev. B **72**, 144408 (2005).

[41] K. H. Hong, A. M. Arevalo-Lopez, M. Coduri, G. M. McNally, and J. P. Attfield, Cation, magnetic, and charge ordering in $MnFe_3O_5$, J. Mater. Chem. C **6**, 3271 (2018).

[42] L. Jiang, S. Yang, M. Zheng, H. Chen, and A. Wu, Synthesis and magnetic properties of nanocrystalline $Gd_3Fe_5O_{12}$ and $GdFeO_3$ powders prepared by sol–gel auto-combustion method, Mater. Res. Bull. **104**, 92 (2018).

[43] M. Mansmann and W. E. Wallace, Magnetic properties of $GdFe_2$ and $DyFe_2$, J. Chem. Phys. **40**, 1167 (1964).

[44] I. W. Modder, H. Bakker, and G. F. Zhou, Defects and magnetism in cubic GdX_2 Laves phase compounds, Phys. B: Condens. Matter **262**, 141 (1999).

[45] M. A. Rafiq, A. Javed, M. N. Rasul, M. Nadeem, F. Iqbal, and A. Hussain, Structural, electronic, magnetic and optical properties of AB_2O_4 (A = Ge, Co and B = Ga, Co) spinel oxides, Mater. Chem. Phys. **257**, 123794 (2021).

[46] A. Tsuzuki, S. Sago, S. I. Hirano, and S. Naka, High temperature and pressure preparation and properties of iron carbides Fe_7C_3 and Fe_3C, J. Mater. Sci. **19**, 2513 (1984).

[47] M. Wakizaka, W.-J. Chun, T. Imaoka, and K. Yamamoto, Synthesis and magnetic properties of sub-nanosized iron carbides on a carbon support, RSC Adv. **12**, 3238 (2022).

[48] D. Williams, P. M. Shand, C. Stark, T. Pekarek, R. Brown, L. Yue, and D. L. Leslie-Pelecky, Curie–Weiss analysis of ferromagnetic and glassy transitions in nanostructured GdAl$_2$, J. Appl. Phys. **93**, 6525 (2003).

[49] T. Yamamoto, K. Kaminaga, D. Saito, D. Oka, and T. Fukumura, Rock salt structure GdO epitaxial thin film with a high ferromagnetic curie temperature, Appl. Phys. Lett. **117**, 052402 (2020).

[50] N. Zhao, W. Wang, X. Lei, Z. Ye, X. Chen, H. Ding, and H. Yang, Synthesis, structure and magnetic properties of Fe$_3$N nanoparticles, J. Mater. Sci.: Mater. Electron. **28**, 15701 (2017).

[51] D. B. McWhan, P. C. Souers, and G. Jura, Magnetic and structural properties of europium metal and europium monoxide at high pressure, Phys. Rev. **143**, 385 (1966).

[52] D. Le Binh, B. J. Ruck, F. Natali, H. Warring, H. J. Trodahl, E.-M. Anton, C. Meyer, L. Ranno, F. Wilhelm, and A. Rogalev, Europium Nitride: A Novel Diluted Magnetic Semiconductor, Phys. Rev. Lett. **111**, 167206 (2013).

[53] V. Pecharsky, K. Gschneidner, S. Y. Dan'kov, and A. Tishin, Magnetocaloric properties of Gd$_3$Al$_2$, in *Cryocoolers 10*, edited by R. G. Ross (Springer, Boston, MA, 2002), pp. 639–645.

[54] E.-M. Anton, J. F. McNulty, B. J. Ruck, M. Suzuki, M. Mizumaki, V. N. Antonov, J. W. Quilty, N. Strickland, and H. J. Trodahl, NdN: An intrinsic ferromagnetic semiconductor, Phys. Rev. B **93**, 064431 (2016).

[55] S. Labroo and N. Ali, Magnetism of rare-earth disilicides, J. Appl. Phys. **67**, 4811 (1998).

[56] X. C. Zhong, M. Zou, H. Zhang, Z. W. Liu, D. C. Zeng, K. A. Gschneidner, and V. K. Pecharsky, Crystal structure and magnetic properties of R$_5$Sn$_4$ alloys, where R is Tb, Dy, Ho, and Er, J. Appl. Phys. **109**, 07A917 (2011).

[57] O. V. Yakubovich, L. V. Shvanskaya, O. V. Dimitrova, O. S. Volkova, and A. N. Vasiliev, Magnetically frustrated synthetic end member Mn$_2$(PO$_4$)OH in the triplite–triploidite family, Dalton Trans. **46**, 8680 (2017).

[58] H. Wada and M. Shiga, Antiferromagnetic properties of EuAl$_2$, Phys. B: Condens. Matter **193**, 25 (1994).

[59] N. V. Tristan, S. A. Nikitin, T. Palewski, K. Skokov, and J. Warchulska, Magnetization of a Gd$_3$Ni single crystal, J. Alloys Compd. **334**, 40 (2002).

[60] J. Roger, V. Babizhetskyy, K. Hiebl, J.-F. Halet, and R. Guérin, Structural chemistry, magnetism and electrical properties of binary Gd silicides and Ho$_3$Si$_4$, J. Alloys Compd. **407**, 25 (2006).

[61] R. J. Pollard, V. H. McCann, and J. B. Ward, Magnetic structures of α-MnS and MnSe from ^{57}Fe Mossbauer spectroscopy, J. Phys. C **16**, 345 (1983).

[62] X. G. Li, M. Sato, S. Takahashi, K. Aoki, and T. Masumoto, Magnetic properties of Gd$_2$Al compound in amorphous, crystalline and nitrided states, J. Magn. Magn. Mater. **212**, 145 (2000).

[63] R. L. Carlin, R. D. Chirico, K. O. Joung, G. K. Shenoy, and D. G. Westlake, Magnetic ordering in GdH$_3$, Phys. Lett. A **75**, 413 (1980).

[64] M. Kolenda, A. Szytula, J. Leciejewicz, A. Pawlukojc, and H. Ptasiewicz-Bak, Magnetic structure of Mn$_3$Ni$_2$Si, J. Magn. Magn. Mater. **89**, 26 (1990).

[65] R. D. Johnson, D. D. Khalyavin, P. Manuel, L. Zhang, K. Yamaura, and A. A. Belik, Magnetic structures of the rare-earth quadruple perovskite manganites RMn$_7$O$_{12}$, Phys. Rev. B **98**, 104423 (2018).

[66] S. Hémon, R. A. Cowley, R. C. C. Ward, M. R. Wells, L. Douysset, and H. Ronnow, Magnetic structure of Gd, GdH$_2$ and NdH$_2$ single crystal films, J. Phys.: Condens. Matter **12**, 5011 (2000).

[67] N. V. Baranov, A. V. Andreev, A. I. Kozlov, G. M. Kvashnin, H. Nakotte, H. Aruga Katori, and T. Goto, Magnetic phase transitions in Gd$_3$Co, J. Alloys Compd. **202**, 215 (1993).

[68] T. Tsutaoka, K. Ueda, and T. Matsushita, Magnetic and electrical properties of Nd$_7$Pt$_3$ studied on single crystals, Phys. B: Condens. Matter **541**, 50 (2018).

[69] S. Ozeki, Y. S. Kwon, Y. Haga, T. Suzuki, G. Kido, and T. Kasuya, de Haas–van Alphen effect in SmSb, Phys. B: Condens. Matter **169**, 499 (1991).

[70] F. P. Missell, R. P. Guertin, and S. Foner, Magnetic properties of NdSb, GdSb, TbSb, DySb, HbSb and ErSb under hydrostatic pressure, Solid State Commun. **23**, 369 (1977).

[71] I. V. Golosovskii and V. P. Plakhty, Magnetic ordering in NdS, Phys. Status Solidi B **56**, 61 (1973).

[72] G. P. Felcher, T. O. Brum, R. J. Gambino, and M. Kuznietz, Magnetic structure of DySb, Phys. Rev. B **8**, 260 (1973).

[73] P. Ĉermák, K. Prokeš, B. Ouladdiaf, M. Boehm, M. Kratochvílová, and P. Javorský, Magnetic structures in the magnetic phase diagram of Ho$_2$RhIn$_8$, Phys. Rev. B **91**, 144404 (2015).

[74] M. Reiffers, S. Ilkovič, B. Idzikowski, J. Šebek, and E. Šantavá, Heat capacity and point-contact spectra of the melt-spun cubic RECu$_5$ compounds (RE – heavy rare earths), J. Phys.: Conf. Ser. **200**, 032061 (2010).

[75] S. Tuncel, J. G. Roquefère, C. Stan, J.-L. Bobet, B. Chevalier, E. Gaudin, R.-D. Hoffmann, U. C. Rodewald, and R. Pöttgen, Rare earth metal rich magnesium compounds RE_4NiMg (RE=Y, Pr–Nd, Sm, Gd–Tm, Lu)—synthesis, structure, and hydrogenation behavior, J. Solid State Chem. **182**, 229 (2009).

[76] C. Wäckerlin, F. Donati, A. Singha, R. Baltic, A.-C. Uldry, B. Delley, S. Rusponi, and J. Dreiser, Strong antiferromagnetic exchange between manganese phthalocyanine and ferromagnetic europium oxide, Chem. Commun. **51**, 12958 (2015).

[77] J. R. Rumble, *CRC Handbook of Chemistry and Physics*, 103rd ed. (CRC Press/Taylor & Francis, London, 2022).

[78] F. Wu, C. Y. Guo, M. Smidman, J. L. Zhang, and H. Q. Yuan, Large magnetoresistance and Fermi surface topology of PrSb, Phys. Rev. B **96**, 125122 (2017).

[79] Y.-Y. Wang, H. Zhang, X.-Q. Lu, L.-L. Sun, S. Xu, Z.-Y. Lu, K. Liu, S. Zhou, and T.-L. Xia, Extremely large magnetoresistance and electronic structure of TmSb, Phys. Rev. B **97**, 085137 (2018).

[80] A. D. Tamerius, S. M. Clarke, M. Gu, J. P. S. Walsh, M. Esters, Y. Meng, C. H. Hendon, J. M. Rondinelli, S. D. Jacobsen, and D. E. Freedman, Discovery of Cu$_3$Pb, Angew. Chem., Int. Ed. **57**, 12809 (2018).

[81] L. Salamakha, E. Bauer, H. Michor, G. Hilscher, H. Müller, R. Svagera, O. Sologub, P. Rogl, J. Hester, T. Roisnel, G. Giester, and S. Mudryi, First-order phase transition in a new CaCu$_5$-related antimonide, CePt$_5$Sb, Chem. Mater. **23**, 4016 (2011).

[82] B. Qian, F. Tang, Y. R. Ruan, Y. Fang, Z. D. Han, X. F. Jiang, J.-M. Zhang, S. Y. Chen, and D. H. Wang, Extremely large magnetoresistance in the nonmagnetic semimetal YbI, J. Mater. Chem. C **6**, 10020 (2018).

[83] L. Heletta and R. Pöttgen, Equiatomic rare earth rhodium plumbides RERhPb (RE=Y, La–Nd, Sm, Gd–Lu) with ZrNiAl-type structure, Z. Naturforsch. **73**, 251 (2018).

[84] E. M. Carnicom, W. W. Xie, E. M. Seibel, T. Klimczuk, and R. J. Cava, The LaPdIn$_4$ indide and elementary properties of the LaTIn$_4$ (T=Ni, Pd, Pt) materials family, J. Alloys Compd. **694**, 682 (2017).

[85] K. Alami-Yadri, D. Jaccard, and P. Link, High-pressure magnetic ordering of YbSi and YbCu$_2$Si$_2$, Phys. B: Condens. Matter **230**, 272 (1997).

[86] U. Wedig, V. Saltykov, J. Nuss, and M. Jansen, Homoatomic stella quadrangula [Tl$_8$]$^{6-}$ in Cs$_{18}$Tl$_8$O$_6$, interplay of spin-orbit coupling, and Jahn-Teller distortion, J. Am. Chem. Soc. **132**, 12458 (2010).

[87] A. Tursina, S. Nesterenko, E. Murashova, Z. Kurenbaeva, Y. Seropegin, H. Noël, T. Roisnel, and D. Kaczorowski, Synthesis, crystal structure and magnetic properties of the new ternary indides REPd$_2$In$_4$ (RE=La, Ce, Pr, Nd), Intermetallics **19**, 1864 (2011).

[88] D. Stoiber, M. Bobnar, P. Höhn, and R. Niewa, Lithium alkaline earth tetrelides of the type Li$_2$$Ae$$Tt$ (Ae=Ca, Ba, Tt=Si, Ge, Sn, Pb): Synthesis, crystal structures and physical properties, Z. Naturforsch. B **72**, 847 (2017).

[89] A. Grykałowska and B. Nowak, Nuclear spin-lattice relaxation in narrow gap semiconductors TiPtSn and ZrPtSn, Intermetallics **15**, 1479 (2007).

[90] R. Paul, P. Sen, and I. Das, Effect of morphology on the magnetic properties of Gd$_2$O$_3$ nanotubes, Phys. E **80**, 149 (2016).

[91] B. Idzikowski, A. Jezierski, and K. Nenkov, Magnetic properties, magnetoresistance, and electronic structure of cubic DyCu$_5$, J. Appl. Phys. **85**, 4744 (1999).

[92] M. Balli, S. Mansouri, D. Z. Dimitrov, P. Fournier, S. Jandl, and J.-Y. Juang, Strong conventional and rotating magnetocaloric effects in TbVo$_4$ crystals over a wide cryogenic temperature range, Phys. Rev. Mater. **4**, 114411 (2020).

[93] M. Balli, S. Jandl, P. Fournier, and M. M. Gospodinov, Anisotropy-enhanced giant reversible rotating magnetocaloric effect in HoMn$_2$O$_5$ single crystals, Appl. Phys. Lett. **104**, 232402 (2014).

[94] M. Balli, S. Jandl, P. Fournier, and D. Z. Dimitrov, Giant rotating magnetocaloric effect at low magnetic fields in multiferroic TbMn$_2$O$_5$ single crystals, Appl. Phys. Lett. **108**, 102401 (2016).

[95] M. Orendáč, S. Gabáni, E. Gažo, G. Pristáš, N. Shitsevalova, K. Siemensmeyer, and K. Flachbart, Rotating magnetocaloric effect and unusual magnetic features in metallic strongly anisotropic geometrically frustrated TmB$_4$, Sci. Rep. **8**, 10933 (2018).

[96] H. Zhang, Y. Li, E. Liu, Y. Ke, J. Jin, Y. Long, and B. Shen, Giant rotating magnetocaloric effect induced by highly texturing in polycrystalline DyNiSi compound, Sci. Rep. **5**, 11929 (2015).

[97] https://docs.materialsproject.org/open-apis/the-materials-api/.

[98] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

[99] K. Lejaeghere, V. V. Speybroeck, G. V. Oost, and S. Cottenier, Error estimates for solid-state density-functional theory predictions: An overview by means of the ground-state elemental crystals, Crit. Rev. Solid State Mater. Sci. **39**, 1 (2014).

[100] Wolverton Research Group, Magpie, https://bitbucket.org/wolverton/magpie/src/master/.

[101] Wikipedia, List of elements by atomic properties (2021), https://en.wikipedia.org/wiki/List_of_elements_by_atomic_properties.

[102] Scikit-learn: Feature importances with a forest of trees, https://scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances.html.