

Compositional optimization of hard-magnetic phases with machine-learning models

Johannes J. Möller,^{1,*} Wolfgang Körner,¹ Georg Krugel,¹ Daniel F. Urban,¹ and Christian Elsässer^{1,2}

¹*Fraunhofer Institute for Mechanics of Materials IWM, Wöhlerstr. 11, 79108 Freiburg, Germany*

²*University of Freiburg, Freiburg Materials Research Center, Stefan-Meier-Str. 21, 79104 Freiburg, Germany*

Machine Learning (ML) plays an increasingly important role in the discovery and design of new materials. In this paper, we demonstrate the potential of ML for materials research using hard-magnetic phases as an illustrative case. We build kernel-based ML models to predict optimal chemical compositions for new permanent magnets, which are key components in many green-energy technologies. The magnetic-property data used for training and testing the ML models are obtained from a combinatorial high-throughput screening based on density-functional theory calculations. Our straightforward choice of describing the different configurations enables the subsequent use of the ML models for compositional optimization and thereby the prediction of promising substitutes of state-of-the-art magnetic materials like $\text{Nd}_2\text{Fe}_{14}\text{B}$ with similar intrinsic hard-magnetic properties but a lower amount of critical rare-earth elements.

1. INTRODUCTION

Machine Learning (ML) is currently establishing itself as an important tool in materials research. ML models have been successfully used for predicting grain boundary energies [1, 2] and mobilities [3] in pure metals, ferroelectric Curie temperatures and classes of perovskites [4–6], and electronic [7, 8] as well as magnetic properties [9–11]. The main applications of ML have, however, been limited to learning properties that are directly related to total/atomic energies or atomic forces [12–26]. We believe that this limitation originates mainly from the availability of high-accuracy density-functional theory (DFT) data for the energies of millions of structures and compositions in internet databases, such as the Materials Project [27], the Open Quantum Materials Database (OQMD) [28, 29], or the Novel Materials Discovery (NOMAD) repository [30]. The potential of ML to accurately predict and to optimize other materials properties, which are less well documented in internet databases, needs thus to be further explored.

The search for new hard-magnetic phases is an ideal application case for demonstrating the benefits of ML in materials research. On the one hand, many hard-magnetic materials are intermetallic phases that contain a multitude of atoms per unit cell. The unit cell of the prominent $\text{Nd}_2\text{Fe}_{14}\text{B}$ phase, for instance, contains 68 atoms. This makes it infeasible even for combinatorial high-throughput screening (HTS) to cover the entire variety of possible phase compositions. For a reasonable variety of screened elements, the number of possible compositions and atomic arrangements within the unit cell easily adds up to billions or even trillions. ML overcomes this limitation by building a *continuous* model for the learned properties based on the training data. This enables a reliable interpolation between the training data and an accurate prediction of properties for unknown compositions. Today’s best hard-magnetic materials, which

are key components in many green-energy technologies, contain substantial amounts of supply-critical rare-earth (RE) elements, such as Dy, Nd, or Sm. There is thus a considerable industrial demand for new materials with good hard-magnetic properties but less RE contents.

The objective of this paper is therefore to construct ML models for hard-magnetic properties and to subsequently use them for compositional optimization in order to identify those chemical compositions that exhibit good magnetic properties, but contain only few RE elements. For training the ML models, we select the ThMn_{12} -type crystal structure [31] from our materials database of hard-magnetic phases [32] as a promising substitute for the RE-rich state-of-the-art materials $\text{Nd}_2\text{Fe}_{14}\text{B}$ and SmCo_5 . Instead of using all the data in our database for training, we use a subset with equally-distributed and equally-spaced supporting points where only one alloying element (besides Fe) is considered. We believe that such a scenario where one holds a data set with not much more than 200 different configurations is representative for a typical systematic density-functional-theory (DFT) study of a certain non-standard property (which in our case is the magneto-crystalline anisotropy).

The paper is organized as follows: In section 2 we describe our computational approach, i.e., how we train the ML models to the hard-magnetic properties of the intermetallic phases in our database. In section 3 we present the results for validating, testing, and optimizing the ML models. These results are discussed in section 4 with focus on the accuracy and reliability of the ML predictions and on the potential of the optimized compositions. The discussion also includes comments on the achievements and limitations of our approach, on possible pathways for future extensions, and on discrepancies between the DFT training data and experimental results. The paper is summarized in section 5.

* Corresponding author: johannes.moeller@iwm.fraunhofer.de

2. METHODOLOGY

2.1. Machine Learning

The list of machine-learning (ML) models that have been used for predicting properties and behaviors of materials includes a variety of methods such as artificial neural networks [34–37], genetic algorithms [33, 34], Gaussian processes [17, 37], decision-tree and random-forest models [5, 36], and kernel-based methods [13, 14, 26, 36]. Which of the different methods is best suited for a certain analysis usually depends on the type (e.g., scalar values or images) and amount (hundreds, thousands or millions of samples) of the available training data and the targeted application of the ML model (value prediction/regression or classification).

In this work, we make use of the kernel-based Support Vector Regression (SVR) method to construct numerical ML models for K_1 , $\mu_0 M$ and E_f (together denoted as target properties y in the following). SVR is a non-linear regression analysis that makes use of the so-called kernel trick. In kernelized ML methods, the kernel function maps the input space, in which the target property y is usually not a linear function of the feature vector \mathbf{x} , into a higher dimensional space where such a linear relationship may exist. SVR is based on the concept of Support Vector Machines (SVMs), which were originally developed in the 1960s [38] for classification purposes. In SVR models, the predicted value y^m (superscript 'm' for model) for a feature vector \mathbf{x} is determined as

$$y^m(\mathbf{x}) = \sum_i w_i k(\mathbf{x}_i, \mathbf{x}) + b, \quad (1)$$

where w_i are the individual weights for each training vector \mathbf{x}_i , $k(\mathbf{x}_i, \mathbf{x})$ is the kernel function, and b is the constant intercept ('b' stands for bias).

The kernel $k(\mathbf{x}_i, \mathbf{x})$ is typically of linear, polynomial or Gaussian (radial basis function, 'rbf') type [39]:

$$k^{\text{linear}}(\mathbf{x}_i, \mathbf{x}) = \mathbf{x}_i \cdot \mathbf{x}, \quad (2)$$

$$k^{\text{poly}}(\mathbf{x}_i, \mathbf{x}) = [\gamma(\mathbf{x}_i \cdot \mathbf{x})]^d, \quad (3)$$

$$k^{\text{rbf}}(\mathbf{x}_i, \mathbf{x}) = \exp[-\gamma \|\mathbf{x}_i - \mathbf{x}\|_2^2]. \quad (4)$$

Here, the subscript '2' indicates the Euclidean (L_2) distance and ' \cdot ' the inner product, d is the degree of the polynomial kernel, and γ represents the width of the respective kernel function.

SVR uses an ε -insensitive loss function [40, 41], which means that optimizing a SVR model involves minimizing the weights w_i . This procedure leads to a flat evolution of $y(\mathbf{x})$ and inherently reduces the risk of over-fitting [42] by simultaneously allowing for some larger deviations (up to a value of ε) between y and y^m for individual samples ('outliers'). In this paper, we use $\varepsilon = 0.1$ T, 1.0 MJ/m³, and 0.01 eV/atom for the ML models for $\mu_0 M$, K_1 , and E_f , respectively. This choice can be considered as the targeted accuracy of the constructed models.

Besides the choice of the kernel function (and its width γ), SVR models have an additional regularization parameter C , which determines the trade-off between the flatness of $y^m(\mathbf{x})$ and the amount up to which deviations larger than ε are tolerated [41] (a higher C allows higher possible values for the weights w_i). As mentioned before, γ determines the width of the kernel and therefore the contribution of neighboring feature vectors to the prediction (a smaller γ resulting in a higher contribution of samples close by). For a given kernel function, both hyperparameters control the complexity of the model and need to be optimized in order to prevent under- and over-fitting. To yield a model that is as universal as possible, both C and γ should be as small as possible. Note that γ has no meaning for a linear kernel.

Extensive descriptions of the SVR method are given in Refs. [40, 41]. The SVR algorithm is used in its implementations in the Python ML package `scikit-learn` [39, 43]. To assess the potential benefit of ML over classical fitting methods, we also parameterized linear regression (LR) functions ($y^m = \sum_j w_j x_j + b$, $j = 0 \dots N_{\text{features}}$) for our data. Note that in this case, contrary to the SVR method described above, the weights w_j are determined for each component of \mathbf{x} , i.e., for the feature x_j , not for each training sample \mathbf{x}_i .

2.2. Material Database

The material database used for training and validating the ML models originates from our previous study on rare-earth-lean intermetallic REA₁₂X compounds [32]. Here, RE is either Ce or Nd, and A can be one of the magnetic transition-metal elements Mn, Fe, Co, and Ni, the non-magnetic elements Ti, V, Cr, Cu, Zn, Al, Si, and P, or a mixture of them. The element X is one of the small interstitial elements B, C, or N.

The REA₁₂X structure is based on the ThMn₁₂ type, which was first reported by de Mooij *et al.* [31]. The symmetry-equivalent Wyckoff sites of this body-centered tetragonal structure (space group #139, I4/mmm symmetry) are the 2a (occupied by RE elements), 8i, 8j, 8f (occupied by A elements), and 2b (occupied by X elements, if any), see Fig. 1 for illustration. Note that the number of atoms per formula unit (14) is only half of all atoms per conventional tetragonal unit cell (28). The tetragonal lattice parameters $a = 8.566$ Å and $c = 4.802$ Å, as well as the crystal coordinates of the Wyckoff sites were taken from the work of Isnard *et al.* [44].

To create the magnetic-properties database, we employ a HTS approach with DFT calculations using the tight-binding linear-muffin-tin-orbital atomic-sphere-approximation method (TB-LMTO-ASA) [45] based on the work of Fähnle *et al.* [46, 47]. The several thousand phases were calculated with the HTS setup described by Drebov *et al.* [48] which allows a fully automated generation of new phases by combinatorial substitution of sets of

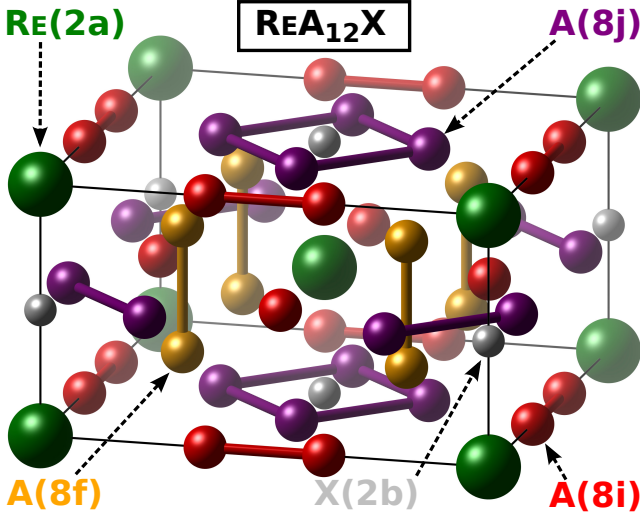


FIG. 1. Crystal structure of the REA_{12}X structure with indicated Wyckoff positions. The bonds between symmetry-equivalent A atoms are drawn for better visualization.

atoms sitting on specific positions. In this study, for each combination of RE and X, all A positions (see Fig. 1) are initially occupied with Fe atoms. All Fe atoms sitting on a specific Wyckoff site are then substituted by other alloying elements A and this substitution is carried out for all combinations of Wyckoff sites. To indicate this substitution schema, we refer to the generated compounds as $\text{RE}(\text{Fe},\text{A})_{12}\text{X}$ in the following. Moreover, we have generated a substantial number of configurations where two alloying elements A and A' are considered and also allow for partial substitution on Wyckoff sites.

By means of DFT calculations, we determine (among other properties) the magnetization $\mu_0 M$, the local magnetic moments, the uniaxial magneto-crystalline anisotropy constant K_1 , and the relative phase-stability energy E_f . Here, the latter is evaluated by comparing the total energy of a compound to the total energies of the elemental constituents (and not with all competing binary or ternary phases). Thereby, its computation is compatible with the approach of a high-throughput analysis and it serves as a valuable first-order estimate for the expectable thermodynamic stability.

The technologically interesting figures of merit of hard-magnetic materials, i.e., the maximal energy product $(BH)_{\max}$ and the anisotropy field H_a are estimated from $\mu_0 M$ and K_1 as follows [49]:

$$(BH)_{\max}^{\text{EST}} = \frac{(0.9\mu_0 M)^2}{4\mu_0}, \quad (5)$$

$$H_a = \frac{2K_1}{\mu_0 M}. \quad (6)$$

Note that Eq. (5) implies the common assumption that ideally about 10% of a processed bulk hard magnet consists of non-magnetic phases [49]. For further details on the HTS approach using TB-LMTO-ASA calculations,

we refer to Refs. [32, 48].

In order to convert our zero-temperature single-crystal, single-domain DFT-HTS results for K_1 and H_a to room-temperature estimates, which can serve as guidelines for experimental efforts, the calculated values need to be divided by four for Nd. For Ce, a division by 35 provides a conservative estimate. These heuristic factors have been derived from benchmark calculations of well-known hard-magnetic materials (see Table II and discussion in Ref. [32]).

For the optimization of the hyperparameters of our ML model (see Secs. 2.1 and 3.1) we used the $\text{REFe}_{12-4z}\text{A}_{4z}\text{X}$ ($z = 0 \dots 3$) subset of all the different $\text{RE}(\text{Fe},\text{A})_{12}\text{X}$ configurations in our database. In this data set, each A position (8i, 8j, 8f) which initially has been occupied by Fe atoms, was subsequently fully replaced by other alloying elements A. For each choice of the $N_X = 3$ choices for X, there exists one configuration without alloying elements (i.e. $z = 0$) and $N_{\text{sub}} = 7$ possible configurations for each of the $N_A = 11$ alloying elements A, namely three for both $z = 1$ and $z = 2$ and one for $z = 3$. For RE = Nd, the data set therefore consists of 234 $[= N_X \cdot (1 + N_{\text{sub}} \times N_A) = 3 \cdot (1 + 7 \times 11)]$ configurations. For RE = Ce, two compounds could not be converged in the given crystal structure, thus the data set contains only 232 configurations.

In total, our database contains 3,080 entries. Besides the 466 configurations used for training the ML models, the data includes 2,614 compounds with multiple substitutional alloying elements and partially chemically heterogeneous occupations of Wyckoff sites. For example the 8i position can be occupied by two Co atoms and six Fe atoms. This remaining data set was used for the subsequent testing (see Sec. 3.2) of our model with the previously optimized hyperparameters. Note that the test data are inherently difficult to be predicted with our realistic choice of training data as they contain lots of situations about which the trained model was not informed at all (i.e., two different alloying elements in the same configuration). The objective of our work is not to create the best possible model (for which we should obviously use the whole data set for training) but to reflect a realistic scenario and therefore use the typical outcome of a systematic DFT study in terms of number and distribution of the data points.

2.3. Material Representation

For a straightforward decoding of the structural and chemical information stored in the feature vectors \mathbf{x} , the construction of \mathbf{x} needs to be bijective. This means that a reversible mapping exists between \mathbf{x} and the represented atomistic configuration. We introduce here a descriptor that sufficiently fulfills this requirement. It contains the number of atoms $n_j^{(i)}$ of each chemical species i resolved

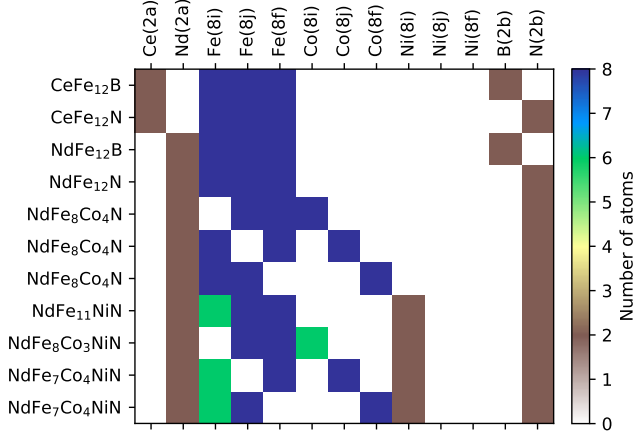


FIG. 2. Illustration of descriptors for selected $\text{RE}(\text{Fe},\text{A})_{12}\text{X}$ compounds. Note that the first seven configurations are included in the training set ($\text{REFe}_{12-4z}\text{A}_{4z}\text{X}$, $z = 0 \dots 4$) whereas the last four compositions belong to the test set (see text for details).

on each of the Wyckoff positions j ,

$$\mathbf{x} = \langle n_1^{(1)}, \dots, n_P^{(1)}, n_1^{(2)}, \dots, n_P^{(2)}, \dots, n_1^{(N)}, \dots, n_P^{(N)} \rangle, \quad (7)$$

where N is the number of different elements in the data set and P is the number of different Wyckoff positions in the crystal structure. When certain positions are occupied only by certain elements (e.g., the 2a site of the REA_{12}X structure is occupied only by RE atoms) the general form of \mathbf{x} given in Eq. 7 can be simplified by leaving out those entries for which $n_j^{(i)}$ is zero throughout the whole data set.

The descriptors for a data set of the REA_{12}X structure with, for instance, two different RE elements on the 2a position ($j = 1$), three different A elements on the 8i, 8j, and 8f positions ($j = 2, 3, 4$), and two different interstitial elements on the 2b position ($j = 5$) are then written as

$$\mathbf{x} = \langle n_1^{(1)}, n_1^{(2)}, n_2^{(3)}, n_3^{(3)}, n_4^{(3)}, n_2^{(4)}, \dots, n_4^{(N-2)}, n_5^{(N-1)}, n_5^{(N)} \rangle \quad (8)$$

and have 13 components in total [instead of $N \cdot P = 7 \cdot 5 = 35$ with Eq. (7)]. Fig. 2 illustrates the descriptors constructed in this way for selected $\text{RE}(\text{Fe},\text{A})_{12}\text{X}$ compounds.

This description has the advantage that structural and chemical information are encoded separately. We briefly note that this representation is somewhat similar to the descriptor which has been used very recently by Faber *et al.* for predicting the formation energies of elpasolite crystals [14].

The construction of this descriptor involves a space group analysis of the crystal structure, for which we use the FINDSYM tool of Stokes & Hatch as implemented in the ISOTROPY Software Suite [50, 51].

2.4. Optimization of Compositions

To efficiently search for compositions that exhibit optimal combinations of $\mu_0 M$ and K_1 , we optimize the model functions $y^m(\mathbf{x})$ (with y being $\mu_0 M$ or K_1) using the stochastic basin-hopping optimization (BHO) algorithm as implemented in the Python package `scipy` [52]. Details of the optimization procedure are given in the A. BHO has the advantage of allowing both *bounded* and *constrained* components of \mathbf{x} . The components of the Wyckoff descriptor have to be *bounded* because each of them has a lower limit (0) and an upper limit (maximal number of atoms per Wyckoff position, e.g. eight atoms for the 8i sites). At the same time, the components of \mathbf{x} are *constrained* because each non-zero component decreases the possible maximal value of all other components representing the same Wyckoff position.

After optimization, the descriptor \mathbf{x}' that fulfills the search criteria, i.e., $y^m(\mathbf{x}') > y_t$ (subscript 't' for target), can be directly mapped to the corresponding chemical composition and distribution of the elements on the Wyckoff positions. This fast and efficient procedure directly benefits from the bijective nature of the Wyckoff descriptor, which allows a reversible mapping between \mathbf{x} and the respective material.

In the optimization, we effectively seek descriptors that maximize the anisotropy coefficient K_1 (and thereby the anisotropy field H_a) by using a relatively conservative target value for $\mu_0 M_t = 1.4$ T since it is well known that the highest values for $\mu_0 M$ are given by compositions containing mainly Fe with (small) contents of Co (see Slater-Pauling curve [53–55]). Note that we are not aiming at minimizing E_f as a consequence of limitations in the input data since our reference database for determining E_f contains only elemental crystal phases to compare with. For this reason, we only use the criterion $E_f^m < E_{ft}$ with a moderate target value $E_{ft} = 0.1$ eV/atom for the optimized compositions (because ± 0.1 eV/atom is about the predictive power of the TB-LMTO-ASA method for compound formation energies).

3. RESULTS

Since the hyperparameters for the SVR algorithm (kernel function k , regularization parameter C , and width γ) are not known beforehand, we first determine the optimal set of hyperparameters that maximizes the predictive power of the model and minimizes its tendency for overfitting. The performance of the optimized ML models is then tested for unseen compositions from our database. Finally, we use an optimization procedure to find the material composition or—more precisely—its descriptor, for which our two figures of merit are within the desired range of values: $\mu_0 M > 1.4$ T and K_1 maximal.

3.1. Training the ML Models

To find the optimal set of hyperparameters for the ML models, we varied the regularization parameter C for the SVR as well as the kernel width γ for various kernel functions. C was varied between 0.1 and 1000 and γ between 10^{-7} and 1.0, with the kernels being linear, polynomial (degree $d = 2$), and rbf functions. The numeric hyperparameters C and γ were varied on a logarithmic grid.

To validate the ML models built with a specific set of hyperparameters, we determine the Pearson correlation coefficient ρ and the mean absolute error (MAE) [42, 56],

$$\rho(y, y^m) = \frac{\text{cov}(y, y^m)}{\sigma_y \sigma_{y^m}}, \quad (9)$$

$$\text{MAE}(y, y^m) = \frac{\sum_{i=1}^N \|y^m - y\|}{N_{\text{samples}}}, \quad (10)$$

where cov is the covariance function, σ denotes the standard deviation, and N_{samples} is the number of samples. The correlation coefficient ρ measures the linear relationship between two (normally distributed) datasets. It varies between -1 and $+1$, where values close to zero indicate that no correlation exists. The limiting values $\rho = +1$ (-1) imply a perfectly direct (indirect) linear proportionality between y^m and y . The MAE (in units of the predicted property) indicates the average difference between y^m and y and should therefore be as small as possible.

To estimate whether the model is over-fitted, i.e., it exactly matches the training samples but nothing else, we performed tenfold cross validation (CV) [42]: The data set used for building and optimizing the ML model consists of 234 (232) samples of type $\text{REFe}_{12-4z}\text{A}_{4z}\text{X}$ for $\text{RE} = \text{Nd}$ (Ce). It is randomly divided into ten subsets of which nine are used to train a new ML model and the tenth is used for validation. This procedure is repeated ten times and the average values of the correlation metrics ρ and MAE for the ten validation sets are used as a measure of the performance of the model with respect to unseen test data.

TABLE I. Optimal set of hyperparameters [kernel k , degree d for polynomial (p) kernels, regularization parameter C , and kernel width γ] and correlation metrics (ρ and MAE) from tenfold CV. $\mu_0 M$ varies between 0 and 2 T; E_f varies between -1 and 1 eV/atom; K_1 varies between 10 and 90 MJ/m³ for Nd and between 50 and 230 MJ/m³ for Ce. For $\mu_0 M$, only configurations with a significant magnetization of $\mu_0 M > 0.25$ T have been used for training (181 for Nd, 179 for Ce).

| Property (unit) | Hyperparameters $k(d), C, \gamma$ | CV (Nd) ρ (MAE) | CV (Ce) ρ (MAE) |
|----------------------------|--------------------------------------|-------------------------|-------------------------|
| $\mu_0 M$ (T) | $p(1), 100, —$ | 0.95 (0.09) | 0.91 (0.11) |
| K_1 (MJ/m ³) | $p(2), 100, 0.1$ | 0.90 (3.9) | 0.88 (10.3) |
| E_f (eV/atom) | $p(2), 0.1, 0.1$ | 0.95 (0.043) | 0.94 (0.04) |

The optimal set of hyperparameters for each com-

bination of ML model and material property were determined by taking those hyperparameters for which C and γ were minimal and satisfying $\rho > 0.95\rho_{\text{max}}$ (ρ_{max} : maximal correlation coefficient for each model/property combination). This ensured that the models were both as accurate and as universal as possible at the same time. The optimal set of hyperparameters and the corresponding correlation metrics ρ and MAE obtained with these hyperparameters are listed in Tab. I. The correlation coefficients obtained with CV are in the range of $\rho = 0.88 \dots 0.95$ for the models with the optimal set of hyperparameters. The CV results are displayed in Fig. 3 for the $\text{Nd}(\text{Fe}, \text{A})_{12}\text{X}$ compounds.

3.2. Testing of the ML models with unseen compositions

For testing our ML models we use test sets that contain configurations, where the the alloying atoms fractionally occupy the 8i Wyckoff site. The resulting predictions for $\text{Nd}(\text{Fe}, \text{A})_{12}\text{X}$ are presented in Fig. 4(top) for SVR and Fig. 4(bottom) for LR. The determined correlation coefficients ρ and MAEs are summarized in the last column of Tab. I. For K_1 and E_f , ρ is somewhat lower than the CV results while the correlation for $\mu_0 M$ is similar to the one determined by CV.

Most of the outliers in the SVR predictions for $\mu_0 M$, see Fig. 4(top), originate from Mn-containing compounds. This can be attributed to the magnetic behavior of Mn, which can change from ferromagnetic to anti-ferromagnetic magnetism depending on the precise atomic arrangement [57]. For K_1 and E_f , most of the outliers are attributed to phases where the non-ferromagnetic elements Al, P, Si, Ti, and Zn represent more than 66 % of all A elements in the compound. In both cases, the evolution of the calculated properties with the elemental concentrations is less systematic and predictions are inherently difficult. It is important to note that the outliers are, however, uncritical for the final optimization because they do not lead to compositions that meet our search criteria.

The test results give us a good indication on how well the ML models will interpolate between the training data, or in other words, how well the models will predict the whole property space of the REA_{12}X crystal structure (for the trained RE, A, and X elements). Note that the achieved MAEs are of the same magnitude as the desired accuracy ε of the SVR models (see section 2.1).

3.3. Optimal Compositions

A selection of the obtained optimal compositions is presented in Tab. II along with their respective values for $\mu_0 M$ and K_1 and the expected values for $(BH)_{\text{max}}$ and H_a estimated according to Eqs. (5-6).

Results of tenfold cross validation (SVR)

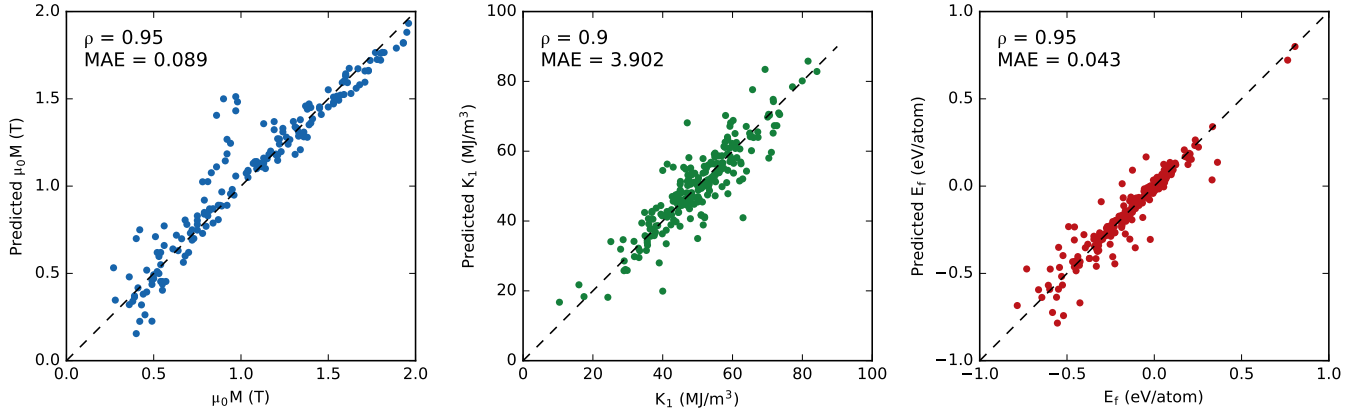
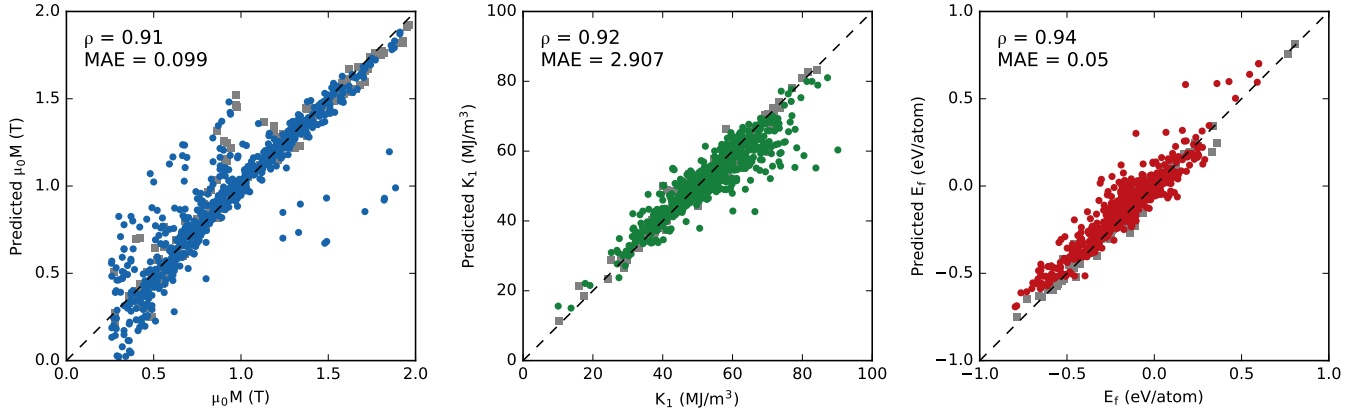


FIG. 3. Results of the tenfold cross validation of the SVR models for $\mu_0 M$ (left), K_1 (center), and E_f (right) for Nd-containing compounds. The colored circles indicate the predictions for each of the ten validation runs. The used (optimal) sets of hyperparameters are given in Tab. I. The MAEs are given in the units of the corresponding axes.

Results of testing with 832 unseen compositions (SVR)



Results of testing with 832 unseen compositions (LR)

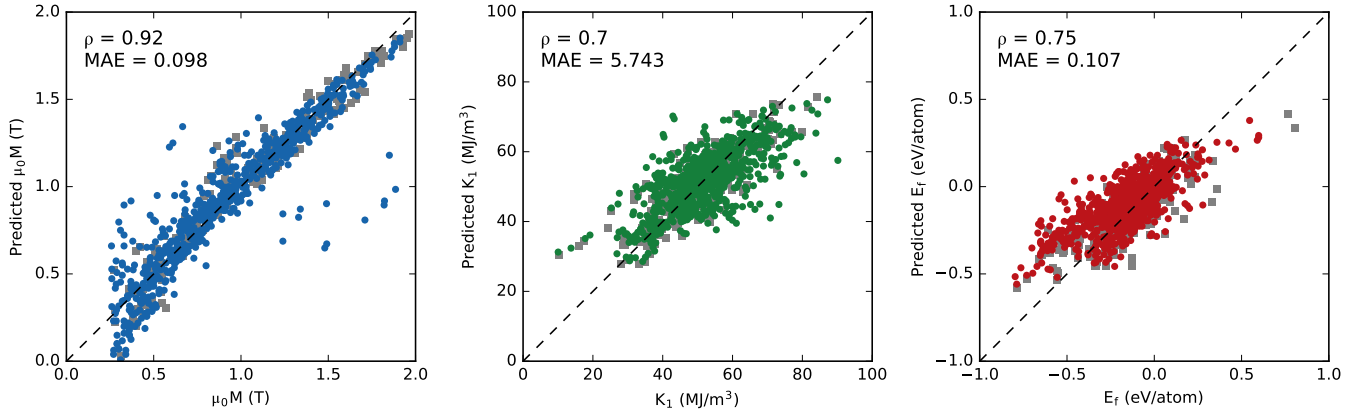


FIG. 4. Testing of the SVR models (top) and LR (bottom) for $\mu_0 M$ (left), K_1 (center), and E_f (right) for Nd-containing compounds. The colored circles indicate the corresponding predictions for the 832 unseen compositions used as test samples. The grey squares are the predictions for the models trained on the whole data set (training and validation data). The used (optimal) sets of hyperparameters for the SVR model are given in Tab. I. The MAEs are given in the units of the corresponding axes.

Even though the test results (section 3.2) already indicate that the ML predictions should be accurate, we veri-

fied the ML-predicted values with *a posteriori* performed TB-LMTO-ASA calculations. In addition to the optimal compositions, the magnetic properties of the most promising candidates drawn from our database [32] as well as the state-of-the-art benchmark materials SmCo_5 and $\text{Nd}_2\text{Fe}_{14}\text{B}$ are presented for completeness. The predicted and calculated values are compared to experimental data whenever available. Note that the values for the formation energy E_f are not given in the table explicitly, but they are negative for all configurations shown.

The optimal compositions given in Tab. II exemplify the versatility of our approach, i.e. the combination of an optimization algorithm with a bijective descriptor. Two compositions were, for instance, optimized under the constraint that they should contain Ti, which is known for its beneficial effect on the phase stability [61]. The predicted $\text{NdFe}_{8.5}\text{Cu}_{3.5}\text{N}$ phase, as another example, avoids the elements Co and Ni, which are both more expensive than Cu.

4. DISCUSSION

In the discussion, we focus on the comparison between the ML predictions and DFT-determined values of the optimal compositions. In addition, we critically assess the potential of the proposed compositions to substitute the state-of-the-art materials NdFe_{14}B and SmCo_5 . Finally, we demonstrate the potential benefit of using kernel-based ML methods over classical LR fitting.

4.1. Comparison between ML predictions and DFT

Although the ML models were trained only to samples containing one alloying element (in addition to Fe), it is remarkable that the predictions for the optimal compositions (which contain often at least two alloying elements) are in overall very good agreement with the DFT calculations performed *a posteriori* (the average ML-DFT difference is below 4 %). A plausible reason for this is the property of iron, other transition metals (TM), and their alloys (whose electronic band structures are mainly formed by the TM d-orbitals) that along the TM series in the periodic systems the band structures and densities of states for a specific crystal structure are only scaled in width but hardly changed in shape with varying number of electrons occupying the bands (see, e.g. Refs. [62, 63]).

The apparent success of the machine to learn the underlying physics should not distract us from the fact that the interpolation behavior between training samples is mainly determined by the kernel function. This property of ML models is visualized in Fig. 5 where the predicted K_1 values for different contents of Co on the 8i site, $n_{8i}(\text{Co}) = 0 \dots 8$ ($= z$), in the $\text{NdFe}_{12-z/2}\text{Co}_{z/2}\text{N}$ phase space are plotted for different kernel functions in comparison with TB-LMTO-ASA results. Whereas with the polynomial kernel ($d = 2$), the nearly linear increase

of K_1 is very well reproduced, the linear kernel fails already in predicting the training data at $z = 0$ and 8, respectively. Due to its higher flexibility, the 'rbf' kernel exactly reproduces the training data points. The drawback of this flexibility is that the deviation from the DFT results becomes very large in the range of z values.

One of the recipes of success for ML of materials properties is therefore the combination of a descriptor, whose components have an almost linear relationship with the modelled properties, and a comparably inflexible kernel function. Since many physical properties have polynomial relationships of low degrees with the material composition, this approach is expected to work equally well for other materials properties.

At the same time, one limitation of our approach is that the employed descriptor encodes all chemical and structural information implicitly. ML Models built with this descriptor are thus not transferable to other crystal structures. Promising *universal* descriptors, which may overcome this shortcoming, have already been proposed for molecular systems [26, 64] and their suitability for use needs to be tested for periodic crystals in the future [13].

4.2. Potential of the optimized compositions

The hard-magnetic properties of both benchmark materials, which contain comparably high amounts of RE elements, are currently known as the upper limits for $(BH)_{\text{max}}$ (NdFe_{14}B) and H_a (SmCo_5). As to be expected, our candidates do not beat these values, but they represent alternatives with significantly reduced RE contents.

In comparison with the best results from the HTS

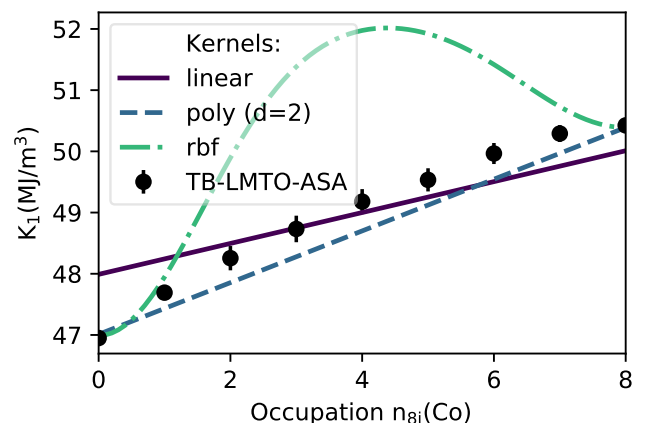


FIG. 5. (Color online) Influence of the kernel function on the prediction of K_1 for unseen compositions $\text{NdFe}_{12-z/2}\text{Co}_{z/2}\text{N}$ ($z = n_{8i}(\text{Co}) = 0 \dots 8$) in comparison to TB-LMTO-ASA results. The model was trained using the linear, polynomial (degree $d = 2$), and 'rbf' kernels. Only the values for $n_{8i}(\text{Co}) = 0$ and 8 were used for training.

TABLE II. Selected compositions with high K_1 and $\mu_0 M$ above 1.4 T found by optimizing the ML models built for the REA₁₂X crystal structure. Predictions of the ML models are indicated with a superscript ‘m’ and compared to *a posteriori* performed DFT (TB-LMTO-ASA) calculations. The Wyckoff positions of Fe substitutes are specified as superscripts (e.g., Co_{0.5}^{1/8i} means that Co occupies one out of eight 8i positions, which is equivalent to 0.5 atoms in the formula unit). The best candidates from the HTS [32] are marked with an asterisk (*). The calculated K_1 values are scaled (after the slash) for better comparability with experiments (see section 2.2 for the heuristic scaling factors). The macroscopic values for $(BH)_{\max}^{\text{EST}}$ and H_a are estimated according to Eqs. (5-6). The available experimental values for $\mu_0 M$, K_1 , $(BH)_{\max}$, and H_a are given in parentheses. E_f is negative for all listed compositions.

| Composition | Predictions of ML Models | | TB-LMTO-ASA Calculations | | Estimations of macroscopic values | |
|---|--------------------------|---------------------------------------|--------------------------|----------------------------|---|-----------------|
| | $\mu_0 M^{\text{m}}$ (T) | K_1^{m} (MJ/m ³) | $\mu_0 M$ (T) | K_1 (MJ/m ³) | $(BH)_{\max}^{\text{EST}}$ (kJ/m ³) | H_a (T) |
| CeFe ₆ Co ₄ ^{8/8f} Cu _{1.5} ^{3/8i} Ti _{0.5} ^{1/8i} N | 1.45 | 165 | 1.45 | 169 / 4.8 | 337 | 233 |
| CeFe ₇ Ni ₄ ^{8/8f} Co ₂ ^{2/8i} N | 1.48 | 167 | 1.57 | 170 / 4.9 | 398 | 216 |
| CeFe ₈ Ni ₄ ^{8/8f} N* | — | — | 1.61 | 167 / 4.8 | 417 | 207 |
| NdFe _{6.5} Co _{3.5} ^{2/8i,5/8f} P _{1.5} ^{3/8j} Ti _{0.5} ^{1/8i} C | 1.42 | 57 | 1.46 | 60 / 15 | 346 | 83 |
| NdFe _{8.5} Cu _{3.5} ^{3/8i,4/8j} N | 1.48 | 57 | 1.57 | 55 / 14 | 397 | 70 |
| NdFe ₈ Ni ₄ ^{8/8f} N* | — | — | 1.68 | 57 / 14 | 457 | 67 |
| <i>Benchmark Materials</i> | | | | | | |
| SmCo ₅ | — | — | 1.07 (1.07 [58]) | 69 (26 [58]) | 184 (219 [59]) | 129 (40.4 [58]) |
| Nd ₂ Fe ₁₄ B | — | — | 1.87 (1.86 [60]) | 19 (4.9 [58]) | 563 (516 [59]) | 20 (6.6 [58]) |

(marked with an asterisk in the table), the K_1 and H_a values of most of the compositions found by optimizing the ML models are similar or higher. The largest improvement is achieved for NdFe_{6.5}Co_{3.5}P_{1.5}Ti_{0.5}C, for which H_a is more than 20 % higher than for the best candidate from the HTS in Ref. [32]. In the future, it would also be interesting to compare our best candidates obtained with BHO to predictions using multi-objective optimization, see Ref. [34] for an example in materials processing.

Finally, intermetallic phases with hypothetically good hard-magnetic properties can exist in reality only if they have sufficient thermodynamic stability. As stated above, in our DFT database of Ref. [32] on which our present ML work is based, the formation energies of intermetallic phases are calculated with the TB-LMTO-ASA method and only with respect to their constituting single-component phases. The ASA is a well suited and justified approximation for calculating hard-magnetic key quantities $\mu_0 M$ and K_1 of REA₁₂-based phases with approximately close-packed crystal structure [32, 47]. For estimating formation-energy differences between competing multi-component phases more precisely, total-energy calculations with more accurate *full-potential* DFT methods are necessary. However, theoretically predicted phases with promising hard-magnetic properties are already valuable information for experimentalists who want to produce hard-magnetic materials by metallurgical means even without prior theoretical stability prediction. Future work on extending DFT databases to hard-magnetic phases with accurate phase-stability information is a desirable task. However, for the present ML work it is not an issue because the presented ML approach to predict chemical compositions of the specific REA₁₂ structure type with good hard-magnetic properties, is not limited

by a limited precision of the thermodynamic-stability data. Once that the DFT database is extended with more accurate *full-potential* DFT data for formation energies of competing multi-component phases, the presented ML approach will be immediately applicable to such an improved database.

4.3. Comparison between ML and LR

As it can be seen from the testing results in Fig. 4, the $\mu_0 M$ correlation metrics for the SVR ML model (with a linear kernel function) are nearly identical to those of LR. This clearly illustrates that the benefits of ML come only into play if a non-linear kernel (as for K_1 and E_f) is used. In our work, this improvement is a factor of 2 for the MAE and around 0.2 for the correlation coefficient.

5. SUMMARY

In this work, we successfully applied a ML approach to a common materials science problem, namely the search for a material composition that optimizes a certain physical property. Based on our HTS database, we trained ML models that reliably predict the hard-magnetic key properties of RE(Fe,A)₁₂X compounds with up to three different non-ferrous alloying elements A based on input data with compounds containing only one alloying element. By optimizing the ML model functions we identified optimal compositions for the REA₁₂X structure that exhibit an increase in magneto-crystalline anisotropy field H_a by more than 10 % as compared to our previous HTS results. Although in both $(BH)_{\max}^{\text{EST}}$ and H_a the predicted

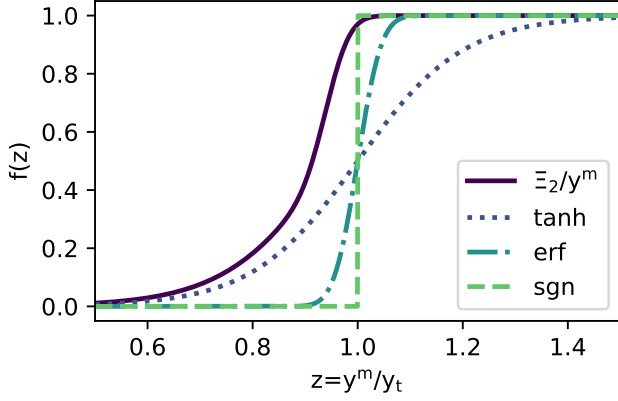


FIG. 6. (Color online) Qualitative comparison of the asymmetric, smooth threshold function $\Xi_2(z)/y^m$ with the point-symmetric tanh, erf, and sgn functions where $z = y^m/y_t$. Note that the arguments and parameters of the tanh, erf, and sgn functions were adjusted here to the threshold value 1 and the range of function values between 0 and 1 in order to be directly comparable with Ξ_2 ; furthermore, for the tanh and erf functions the slope was adjusted in a similar manner as for Ξ_2 , cf. Eq. (A3).

optimal compositions do not excel the current state-of-the-art materials SmCo_5 and $\text{Nd}_2\text{Fe}_{14}\text{B}$ at the same time, they can still be valuable for bridging the gap between these [comparably low $(BH)_{\text{max}}^{\text{EST}}$ and high H_a for SmCo_5 and vice versa for $\text{Nd}_2\text{Fe}_{14}\text{B}$] while having significantly lower RE contents and being therefore more cost efficient and less supply critical.

Our results clearly highlight the potential of ML methods for materials discovery and design. With an appropriate choice of model hyperparameters and material representation it is possible to accurately predict the materials properties in the entire compound space and to identify compositions that optimize the learned property. In the end, the applicability of ML is only limited by the availability and accuracy of training data. For the future, we believe that ML of materials properties will be especially helpful, when the underlying structure-composition-property relationships are not yet completely clarified, as e.g., in the case of K_1 for novel RE-lean magnets.

6. ACKNOWLEDGMENTS

Financial support for this work was provided by the Fraunhofer Lighthouse Project *Critical Rare Earths*.

Appendix A: Optimization details

Using basin-hopping optimization, we intend to determine the descriptor \mathbf{x} that minimizes the following function:

$$F(\mathbf{x}) = -\Xi_1(E_f^m, E_t)\Xi_2(\mu_0 M^m, \mu_0 M_t)\Xi_2(K_1^m, K_{1t}) \quad (\text{A1})$$

where the subscript 't' denotes the desired target values. Note that $\mu_0 M^m$, K_1^m , and E_f^m are functions of \mathbf{x} , which is omitted in Eq. (A1) for better readability. The smooth threshold function Ξ_1 (Ξ_2) becomes one ($\mu_0 M^m$, K_1^m) below (above) the target value E_{ft} ($\mu_0 M_t$, K_{1t}). Both functions converge to zero otherwise. The functions Ξ_1 (for E_f) and Ξ_2 (for $\mu_0 M$ for K_1) are defined as follows:

$$\Xi_1(z) = \frac{1}{2}\text{erfc}\left[5\left(z - \frac{1}{2}\right)\right] \quad (\text{A2})$$

$$\Xi_2(z)/y^m = 1 - \frac{1}{4}\{\tanh[-5(z - 0.95)] + 1\} \{\text{erf}[-20(z - 0.95)] + 1\} \quad (\text{A3})$$

with $z = y^m/y_t$ and y being either $\mu_0 M$, K_1 , or E_f , erf is the error function and erfc is its complement. Obviously, the complementary error function in $\Xi_1(z)$ appears as being a logical choice for a function that should penalize values which become higher than certain target values.

The functional form of $\Xi_2(z)$, on the other hand, needs some more explanation. It was chosen such that values smaller than the target value are less strongly penalized than values above it. By comparing Ξ_2 with its two main point-symmetric constituents, namely the tanh and erf functions, this behavior is visually exemplified in Fig. 6. The somewhat unorthodox choice of $\Xi_2(z)$ has the objective to 'guide' the optimization algorithm in the correct direction (increasing $\mu_0 M$ and K_1 values). Above the thresholds, however, all values contribute equally to $F(\mathbf{x})$, i.e., only by their magnitudes.

[1] Kiyohara, S., Oda, H., Miyata, T., Mizoguchi, T.. Prediction of interface structures and energies via virtual screening. *Sci Adv* 2016;2(11).
[2] Kiyohara, S., Oda, H., Tsuda, K., Mizoguchi, T.. Acceleration of stable interface structure searching using a kriging approach. *Jpn J Appl Phys* 2016;55(4):045502. URL: <http://stacks.iop.org/1347-4065/55/i=4/a=045502?key=crossref.6edc2feef01670642598a371ff15638a>. doi:doi:10.7567/JJAP.55.045502.

[3] Rosenbrock, C.W., Homer, E.R., Csányi, G., Hart, G.L.W.. Discovering the Building Blocks of Atomic Systems using Machine Learning 2017;URL: <http://arxiv.org/abs/1703.06236>. arXiv:1703.06236.
[4] Balachandran, P.V., Xue, D., Lookman, T.. Structure-Curie temperature relationships in BaTiO_3 -based ferroelectric perovskites: Anomalous behavior of $(\text{Ba,Cd})\text{TiO}_3$ from DFT, statistical inference, and experiments. *Phys Rev B* 2016;93(14):1–12. doi:doi:10.1103/PhysRevB.93.144111.

- [5] Pilania, G., Balachandran, P.V., Gubernatis, J.E., Lookman, T.. Classification of ABO₃ perovskite solids: A machine learning study. *Acta Crystallogr Sect B Struct Sci Cryst Eng Mater* 2015;71:507–513. doi:doi:10.1107/S2052520615013979.
- [6] Yang, L., Dacek, S., Ceder, G.. Proposed definition of crystal substructure and substructural similarity. *Phys Rev B* 2014;90(5):054102. URL: <http://link.aps.org/doi/10.1103/PhysRevB.90.054102>. doi:doi:10.1103/PhysRevB.90.054102.
- [7] Lopez-Bezanilla, A., Von Lilienfeld, O.A.. Modeling electronic quantum transport with machine learning. *Phys Rev B - Condens Matter Mater Phys* 2014;89(23):1–5. doi:doi:10.1103/PhysRevB.89.235411. [arXiv:arXiv:1401.8277v1](https://arxiv.org/abs/1401.8277v1).
- [8] Montavon, G., Rupp, M., Gobre, V., Vazquez-Mayagoitia, A., Hansen, K., Tkatchenko, A., et al. Machine learning of molecular electronic properties in chemical compound space. *New J Phys* 2013;15. doi:doi:10.1088/1367-2630/15/9/095003. [arXiv:1305.7074](https://arxiv.org/abs/1305.7074).
- [9] Cheng, W.. Prediction of Magnetic Remanence of Nd-FeB Magnets by using novel machine learning intelligence approach- Support Vector Regression. *Proc 2014 IEEE 13th Int'l Conf Cogn Informatics Cogn Comput* 2014;13:431–435.
- [10] Dam, H.C., Pham, T.L., Ho, T.B., Nguyen, A.T., Nguyen, V.C.. Data mining for materials design: A computational study of single molecule magnet. *J Chem Phys* 2014;140(4):044101. URL: <http://scitation.aip.org/content/aip/journal/jcp/140/4/10.1063/1.4862156>. doi:doi:10.1063/1.4862156.
- [11] Kusne, A.G., Gao, T., Mehta, A., Ke, L., Nguyen, M.C., Ho, K.M., et al. On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets. *Sci Rep* 2015;4(1):6367. URL: <http://www.nature.com/articles/srep06367>. doi:doi:10.1038/srep06367.
- [12] Ghiringhelli, L.M., Vybiral, J., Levchenko, S.V., Draxl, C., Scheffler, M.. Big data of materials science: Critical role of the descriptor. *Phys Rev Lett* 2015;114(10):1–5. doi:doi:10.1103/PhysRevLett.114.105503. [arXiv:arXiv:1411.7437v2](https://arxiv.org/abs/1411.7437v2).
- [13] Faber, F., Lindmaa, A., von Lilienfeld, O.A., Armiento, R.. Crystal structure representations for machine learning models of formation energies. *Int J Quantum Chem* 2015;115(16):1094–1101. URL: <http://doi.wiley.com/10.1002/qua.24917>. doi:doi:10.1002/qua.24917.
- [14] Faber, F.A., Lindmaa, A., von Lilienfeld, O.A., Armiento, R.. Machine Learning Energies of 2 Million Elpasolite (A B C 2 D 6) Crystals. *Phys Rev Lett* 2016;117(13):135502. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.117.135502>. doi:doi:10.1103/PhysRevLett.117.135502.
- [15] Balabin, R.M., Lomakina, E.I.. Support vector machine regression (LS-SVM)an alternative to artificial neural networks (ANNs) for the analysis of quantum chemistry data? *Phys Chem Chem Phys* 2011;13(24):11710. URL: <http://xlink.rsc.org/?DOI=c1cp00051a>. doi:doi:10.1039/c1cp00051a.
- [16] Bartók, A.P., Gillan, M.J., Manby, F.R., Csányi, G.. Machine-learning approach for one- and two-body corrections to density functional theory: Applications to molecular and condensed water. *Phys Rev B* 2013;88(5):054104. URL: <http://link.aps.org/doi/10.1103/PhysRevB.88.054104>. doi:doi:10.1103/PhysRevB.88.054104.
- [17] Bartók, A.P., Payne, M.C., Kondor, R., Csányi, G.. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys Rev Lett* 2010;104(13):136403. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.104.136403>. doi:doi:10.1103/PhysRevLett.104.136403.
- [18] Li, Z., Kermode, J.R., De Vita, A.. Molecular Dynamics with On-the-Fly Machine Learning of Quantum-Mechanical Forces. *Phys Rev Lett* 2015;114(March):096405. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.114.096405>. doi:doi:10.1103/PhysRevLett.114.096405.
- [19] Botu, V., Ramprasad, R.. Learning scheme to predict atomic forces and accelerate materials simulations. *Phys Rev B - Condens Matter Mater Phys* 2015;92(9):1–5. doi:doi:10.1103/PhysRevB.92.094306. [arXiv:1505.02701](https://arxiv.org/abs/1505.02701).
- [20] Botu, V., Ramprasad, R.. Adaptive machine learning framework to accelerate ab initio molecular dynamics. *Int J Quantum Chem* 2015;115(16):1074–1083. URL: <http://doi.wiley.com/10.1002/qua.24836>. doi:doi:10.1002/qua.24836.
- [21] Oganov, A.R., Valle, M.. How to quantify energy landscapes of solids. *J Chem Phys* 2009;130(10):104504. URL: <http://aip.scitation.org/doi/10.1063/1.3079326>. doi:doi:10.1063/1.3079326.
- [22] Suzuki, T., Tamura, R., Miyazaki, T.. Machine learning for atomic forces in a crystalline solid: Transferability to various temperatures. *Int J Quantum Chem* 2017;117(1):33–39. URL: <http://doi.wiley.com/10.1002/qua.25307>. doi:doi:10.1002/qua.25307.
- [23] Snyder, J.C., Rupp, M., Hansen, K., Müller, K.R., Burke, K.. Finding Density Functionals with Machine Learning. *Phys Rev Lett* 2012;108(25):253002. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.108.253002>. doi:doi:10.1103/PhysRevLett.108.253002.
- [24] Pilania, G., Wang, C., Jiang, X., Rajasekaran, S., Ramprasad, R.. Accelerating materials property predictions using machine learning. *Sci Rep* 2013;3:2810. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3786293&tool=pmcentrez&rendertype=abstract>. doi:doi:10.1038/srep02810.
- [25] Huan, T.D., Mannodi-Kanakkithodi, A., Ramprasad, R.. Accelerated materials property predictions and design using motif-based fingerprints. *Phys Rev B - Condens Matter Mater Phys* 2015;92(1):1–10. doi:doi:10.1103/PhysRevB.92.014106. [arXiv:1503.07503v2](https://arxiv.org/abs/1503.07503v2).
- [26] Rupp, M., Tkatchenko, A., Müller, K.R., von Lilienfeld, O.A.. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys Rev Lett* 2012;108(5):058301. URL: [http://eprints.pascal-network.org/archive/00009418/\\$\delimitator\\$026E30F\\$nhhttp://link.aps.org/doi/10.1103/PhysRevLett.108.058301](http://eprints.pascal-network.org/archive/00009418/\delimitator026E30F$nhhttp://link.aps.org/doi/10.1103/PhysRevLett.108.058301). doi:doi:10.1103/PhysRevLett.108.058301. [arXiv:1109.2618](https://arxiv.org/abs/1109.2618).
- [27] Jain, A., Ong, S.P., Hautier, G., Chen, W., Richards, W.D., Dacek, S., et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater* 2013;1(1):011002. URL: <http://aip.scitation.org/doi/10.1063/1.4812323>. doi:doi:10.1063/1.4812323.

- [28] 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 2013;65(11):1501–1509. URL: <http://link.springer.com/10.1007/s11837-013-0755-4>. doi:doi:10.1007/s11837-013-0755-4.
- [29] Kirklin, S., Saal, J.E., Meredig, B., Thompson, A., Doak, J.W., Aykol, M., et al. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. npj Comput Mater 2015;1(October):15010. URL: <http://dx.doi.org/10.1038/npjcompumats.2015.10http://www.nature.com/articles/npjcompumats201510>. doi:doi:10.1038/npjcompumats.2015.10.
- [30] Novel Materials Discovery (NOMAD) Repository, <https://repository.nomad-coe.eu/>. URL: <https://repository.nomad-coe.eu/>.
- [31] De Mooij, D., Buschow, K.. Some novel ternary ThMn12-type compounds. J Less Common Met 1988;136(2):207–215. URL: <http://linkinghub.elsevier.com/retrieve/pii/0022508888904249>. doi:doi:10.1016/0022-5088(88)90424-9.
- [32] Körner, W., Krugel, G., Elsässer, C.. Theoretical screening of intermetallic ThMn12-type phases for new hard-magnetic compounds with low rare earth content. Sci Rep 2016;6(April):24686. URL: <http://www.nature.com/srep/2016/160421/srep24686/full/srep24686.html>. doi:doi:10.1038/srep24686.
- [33] Giri, B.K., Hakanen, J., Miettinen, K., Chakraborti, N.. Genetic programming through bi-objective genetic algorithms with a study of a simulated moving bed process involving multiple objectives. Appl Soft Comput 2013;13(5):2613–2623. URL: <https://www.sciencedirect.com/science/article/pii/S1568494612005091>. doi:doi:10.1016/J.ASOC.2012.11.025.
- [34] Chugh, T., Chakraborti, N., Sindhya, K., Jin, Y.. A data-driven surrogate-assisted evolutionary algorithm applied to a many-objective blast furnace optimization problem. Mater Manuf Process 2017;32(10):1172–1178. URL: <https://www.tandfonline.com/doi/full/10.1080/10426914.2016.1269923>. doi:doi:10.1080/10426914.2016.1269923.
- [35] Behler, J., Parrinello, M.. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. Phys Rev Lett 2007;98(14):146401. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.98.146401>. doi:doi:10.1103/PhysRevLett.98.146401.
- [36] Chowdhury, A., Kautz, E., Yener, B., Lewis, D.. Image driven machine learning methods for microstructure recognition. Comput Mater Sci 2016;123:176–187. URL: <http://dx.doi.org/10.1016/j.commatsci.2016.05.034>. doi:doi:10.1016/j.commatsci.2016.05.034.
- [37] Artrith, N., Urban, A., Ceder, G.. Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species. Phys Rev B 2017;96(1):014112. URL: <http://link.aps.org/doi/10.1103/PhysRevB.96.014112>. doi:doi:10.1103/PhysRevB.96.014112.
- [38] Vapnik, V., Lerner, A.. Pattern recognition using generalized portrait method. Autom Remote Control 1963;24:774–780.
- [39] Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., et al. Scikit-learn: Machine Learning in Python. 2017. URL: <http://scikit-learn.org/stable/modules/metrics.html>.
- [40] Vapnik, V.N.. The nature of statistical learning theory. Springer; 1995. ISBN 0387945598. URL: <http://dl.acm.org/citation.cfm?id=211359>.
- [41] Smola, A., Scholkopf, B.. A Tutorial on Support Vector Regression. Stat Comput 2004;14:199–222. URL: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.41.1452>. doi:doi:10.1023/B:STCO.0000035301.49549.88.
- [42] Witten, I.H., Frank, E., Hall, M.A.. Data mining : practical machine learning tools and techniques. Morgan Kaufmann; 2011. ISBN 9780123748560.
- [43] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res 2011;12(Oct):2825–2830.
- [44] Isnard, O., Miraglia, S., Guillot, M., Fruchart, D.. Hydrogen effects on the magnetic properties of RFe11Ti compounds. J Alloys Compd 1998;275-277:637–641. URL: <http://www.sciencedirect.com/science/article/pii/S09255838898004095>. doi:doi:10.1016/S0925-5838(98)00409-5.
- [45] Andersen, O.K.. Linear methods in band theory. Phys Rev B 1975;12(8):3060–3083. URL: <https://link.aps.org/doi/10.1103/PhysRevB.12.3060>. doi:doi:10.1103/PhysRevB.12.3060.
- [46] Hummler, K., Fähnle, M.. Ab initio calculation of local magnetic moments and the crystal field in $\text{scrR}_2\text{Fe}_{14}\text{B}$ ($\text{scrR} = \text{Gd}, \text{Tb}, \text{Dy}, \text{Ho}, \text{and Er}$). Phys Rev B 1992;45(6):3161–3163. URL: <https://link.aps.org/doi/10.1103/PhysRevB.45.3161>. doi:doi:10.1103/PhysRevB.45.3161.
- [47] Fähnle, M., Hummler, K., Liebs, M., Beuerle, T.. Ab initio electron theory for hard-magnetic rare-earth-transition-metal intermetallics. Appl Phys A Solids Surfaces 1993;57(1):67–76. doi:doi:10.1007/BF00331219.
- [48] Drebov, N., Martinez-Limia, A., Kunz, L., Gola, A., Shigematsu, T., Eckl, T., et al. Ab initio screening methodology applied to the search for new permanent magnetic materials. New J Phys 2013;15(12):125023. URL: <http://stacks.iop.org/1367-2630/15/i=12/a=125023?key=crossref.ef50e177df50dfbd249161985105366a>. doi:doi:10.1088/1367-2630/15/12/125023.
- [49] Hirayama, Y., Takahashi, Y., Hirose, S., Hono, K.. NdFe12Nx hard-magnetic compound with high magnetization and anisotropy field. Scr Mater 2015;95:70–72. URL: <http://www.sciencedirect.com/science/article/pii/S1359646214004163>. doi:doi:10.1016/j.scriptamat.2014.10.016.
- [50] Stokes, H.T., Hatch, D.M.. ISOTROPY Software Suite. 2017. URL: <http://iso.byu.edu>.
- [51] Stokes, H.T., Hatch, D.M.. FINDSYM: program for identifying the space-group symmetry of a crystal. J Appl Crystallogr 2005;38(1):237–238. URL: <http://scripts.iucr.org/cgi-bin/paper?S0021889804031528>. doi:doi:10.1107/S0021889804031528.
- [52] Jones, E., Oliphant, T., Peterson, P.. SciPy: Open source scientific tools for Python. 2017. URL: <http://www.scipy.org/>.
- [53] Slater, J.C.. Electronic Structure of Alloys. J Appl Phys 1937;8(6):385–390. URL: <http://aip.scitation.org/>

- doi:10.1063/1.1710311. doi:doi:10.1063/1.1710311.
- [54] Bardos, D.I.. Mean Magnetic Moments in bcc FeCo Alloys. J Appl Phys 1969;40(3):1371–1372. URL: <http://aip.scitation.org/doi/10.1063/1.1657673>. doi:doi:10.1063/1.1657673.
- [55] Miyake, T., Terakura, K., Harashima, Y., Kino, H., Ishibashi, S.. First-Principles Study of Magnetocrystalline Anisotropy and Magnetization in NdFe 12 , NdFe 11 Ti, and NdFe 11 TiN. J Phys Soc Japan 2014;83(4):043702. URL: <http://journals.jps.jp/doi/pdf/10.7566/JPSJ.83.043702> <http://journals.jps.jp/doi/10.7566/JPSJ.83.043702>. doi:doi:10.7566/JPSJ.83.043702.
- [56] Hastie, T., Tibshirani, R., Friedman, J.H.J.H.. The elements of statistical learning : data mining, inference, and prediction. 2009. ISBN 9780387848570.
- [57] Butcher, T.A., Körner, W., Krugel, G., Elsässer, C.. Dependence of magnetisation and magnetocrystalline anisotropy on site distribution of alloying elements in RE-TM phases with ThMn 12 structure. J Magn Magn Mater 2017;441:1–5. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0304885316327780>. doi:doi:10.1016/j.jmmm.2017.05.011.
- [58] Coey, J.M.D.. Magnetism and Magnetic Materials. 2010. ISBN 0521816149. URL: <http://books.google.com/books?id=Ie72CFd-eSEC&pgis=1>. doi:doi:10.1017/CBO9780511845000. arXiv:arXiv:1011.1669v3.
- [59] Buschow, K.H.J.. Concise encyclopedia of magnetic and superconducting materials. Elsevier; 2005. ISBN 9780080457659. URL: <https://books.google.de/books?hl=en&lr=&id=N9mvytGEBtwC&oi=fnd&pg=PP1&dq=buschow+concise+encyclopedia&ots=NaszWwuSKU&sig=AYFKz1K1XOL6qNERBU5NbPmeGp8{#}v=onepage&q=buschowconciseencyclopedia&f=false>.
- [60] Herbst, J.F.. R2Fe14B materials: Intrinsic properties and technological aspects. Rev Mod Phys 1991;63(4):819–898. doi:doi:10.1103/RevModPhys.63.819.
- [61] Buschow, K.. Permanent magnet materials based on tetragonal rare earth compounds of the type RFe12xMx. J Magn Magn Mater 1991;100(1-3):79–89. URL: <http://linkinghub.elsevier.com/retrieve/pii/030488539190813P>. doi:doi:10.1016/0304-8853(91)90813-P.
- [62] Moruzzi, V.L., Janak, J.F., Williams, A.R.A.R.. Calculated electronic properties of metals. Pergamon Press; 1978. ISBN 9781483152080.
- [63] Pettifor, D.G.. Bonding and Structure of Molecules and Solids. Oxford: Clarendon Press; 1995.
- [64] Von Lilienfeld, O.A., Ramakrishnan, R., Rupp, M., Knoll, A.. Fourier series of atomic radial distribution functions: A molecular fingerprint for machine learning models of quantum chemical properties. Int J Quantum Chem 2015;115(16):1084–1093. doi:doi:10.1002/qua.24912. arXiv:1307.2918.