

Interpretable Machine Learning in Solid-State Chemistry, with Applications to Perovskites, Spinel, and Rare-Earth Intermetallics: Finding Descriptors Using Decision Trees

Balaranjan Selvaratnam, Anton O. Oliynyk, and Arthur Mar*



Cite This: *Inorg. Chem.* 2023, 62, 10865–10875



Read Online

ACCESS |



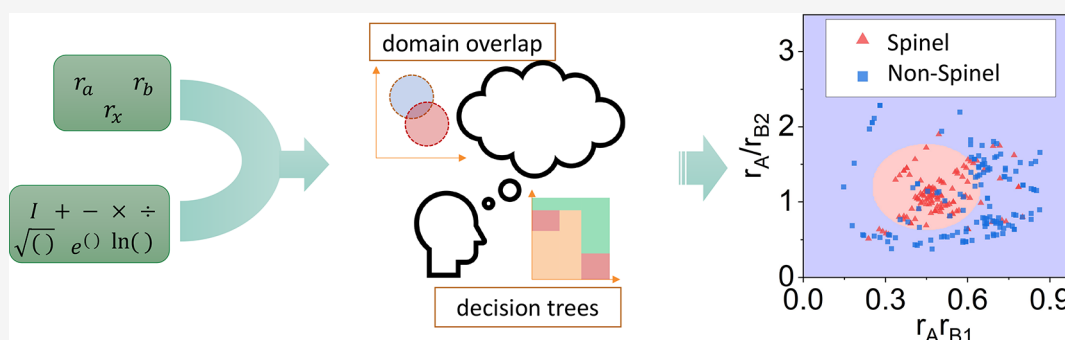
Metrics & More



Article Recommendations



Supporting Information



ABSTRACT: Machine-learning methods have exciting potential to aid materials discovery, but their wider adoption can be hindered by the opaqueness of many models. Even if these models are accurate, the inability to understand the basis for the predictions breeds skepticism. Thus, it is imperative to develop machine-learning models that are explainable and interpretable so that researchers can judge for themselves if the predictions are consistent with their own scientific understanding and chemical insight. In this spirit, the sure independence screening and sparsifying operator (SISSO) method was recently proposed as an effective way to identify the simplest combination of chemical descriptors needed to solve classification and regression problems in materials science. This approach uses domain overlap (DO) as the criterion to find the most informative descriptors in classification problems, but sometimes a low score can be assigned to useful descriptors when there are outliers or when samples belonging to a class are clustered in different regions of the feature space. Here, we present a hypothesis that the performance can be improved by implementing decision trees (DT) instead of DO as the scoring function to find the best descriptors. This modified approach was tested on three important structural classification problems in solid-state chemistry: perovskites, spinels, and rare-earth intermetallics. In all cases, the DT scoring gave better features and significantly improved accuracies of ≥ 0.91 for the training sets and ≥ 0.86 for the test sets.

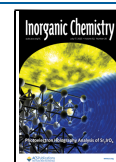
1. INTRODUCTION

Machine-learning methods promise to advance discovery in synthesis, materials, catalysts, and drug design, among many possible areas in chemistry.¹ These methods relate chemical descriptors (the input) to the large knowledge base of structures and properties (the output) of substances. The relationships between the input and output, which are often represented as data in a high-dimensional space, are usually not straightforward. Many machine-learning models are then regarded as “black boxes” if these relationships are too obscure to be understandable by human users (as in neural networks or t-distributed stochastic neighbor embedding) or if too many decision-making entities are involved (as in ensemble trees). Machine-learning models are prone to overfitting, resulting in unnecessarily complicated expressions to arrive at answers that may be correct but for the wrong reasons.² Furthermore, data in materials science are not immune to bias or even

manipulation.^{3–5} The typical quantitative metrics used to assess machine-learning classification models are accuracy [fraction of correct predictions = (true positives + true negatives)/all predictions], precision [fraction of accurate positive predictions = true positives/(true positives + false positives)], and recall [fraction of actual positives accurately predicted = true positives/(true positives + false negatives)]. These metrics only reveal how well a model performs the task based on the outcome rather than how the process works and what factors influence the results.⁶ Similar to other high-stakes

Received: April 10, 2023

Published: June 30, 2023



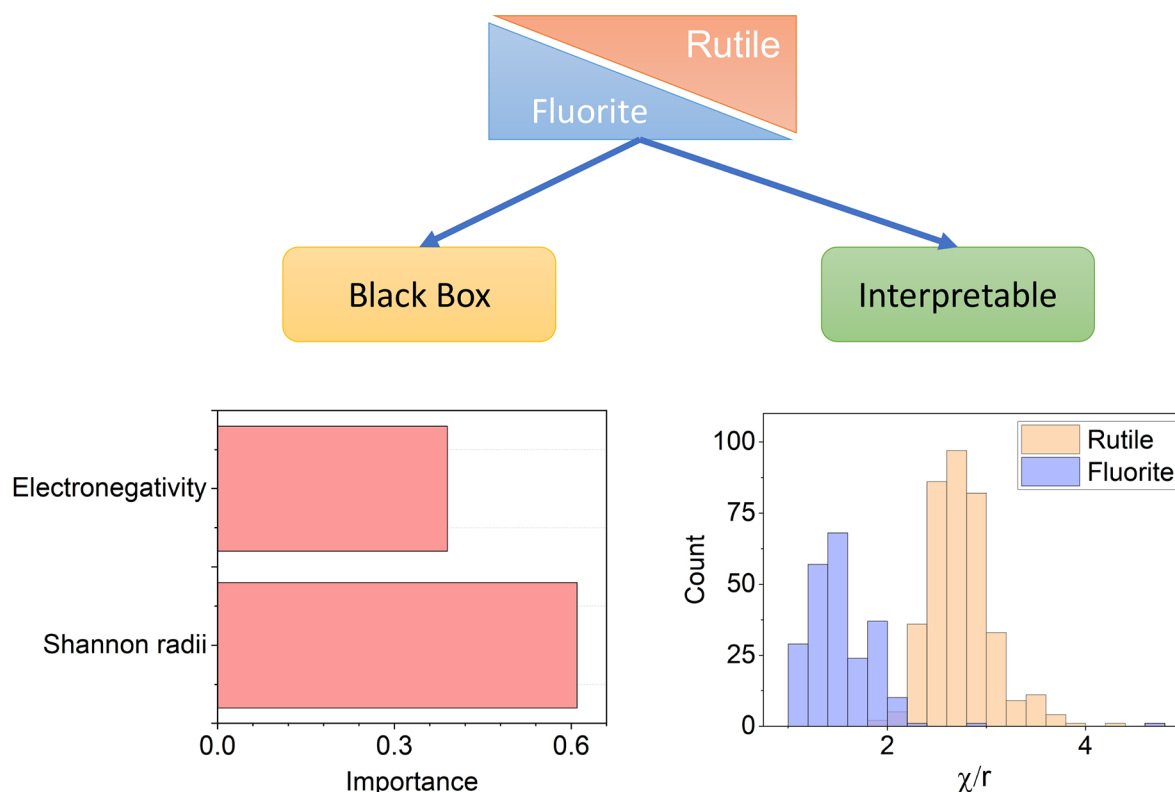


Figure 1. Metal oxides with rutile or fluorite structures can be classified equally well by various machine-learning models, which show that electronegativity and radii are determining factors. The black-box model (based on random forest) identifies the importance of these features but not their relationship. The interpretable model (based on SISSO) shows that the expression χ/r serves as an effective tolerance factor.

applications (such as health care or loan applications), understanding the reasons for a prediction in materials science may be equally or more important than the result.^{7–10} Knowing the relationship between the inputs and outputs enables researchers to judge the quality of predictions, correct for biases, enhance theory, develop general principles, and design further experiments to test the limits of the models.¹¹ The challenge is to seek explanations that balance simplicity with completeness.

Thus, there is impetus to develop machine-learning models that are interpretable, meaning that statements about cause and effect can be formulated, and explainable, meaning that the results can be justified by connecting to selected parameters. The distinction between interpretability and explainability is sometimes confounded.⁹ It is certainly possible for a black-box model to be explainable *post hoc* even if the reasons may not be fully satisfying. There is also no basis for the belief that accuracy must be sacrificed if models are designed to be more interpretable. Model-agnostic metrics, such as permutation importance and Shapley values,^{12,13} do exist that offer explainability (by noting how predictions are affected by changes in feature values) but do not improve interpretability. The features must be carefully considered so that they are informative and meaningful to domain experts.^{14,15}

To illustrate these considerations as applied to solid-state chemistry, consider the problem of determining whether the rutile or fluorite structure is preferred by metal oxides MO_2 .^{16,17} Although simple chemical concepts largely suffice to explain the structural preference, it is interesting to see whether machine-learning methods arrive at the same conclusions. The data set consists of 28 binaries, 256 ternaries,

257 quaternaries, and 55 higher multinaries (see the [Methods](#) section for details).¹⁸ Applying either random forest or sure independence screening and sparsifying operator (SISSO) algorithms results in models that perform equally well ([Figure 1](#)). The random forest model, a black-box one, reveals that the electronegativity χ and radius r of the metal M (or their averages if there are multiple metal atoms) are important features but does not indicate the relationship. The SISSO model, an interpretable one, shows that the ratio of these chemical features acts as a tolerance factor to separate fluorite ($\chi/r \leq 2.2$) and rutile ($\chi/r > 2.2$) structures. This result facilitates the rationalization in chemical terms that the fluorite structure is favored over the rutile structure when the M–O bonding character is more ionic (low χ) and when the metal atoms have large size (high r), consistent with the higher coordination number of 8 in fluorite, which also agrees with the classical radius ratio rules of Pauling.¹⁹

Many questions in solid-state chemistry can be formulated in a manner similar to classification problems. What is the structure type of a given compound (rutile or fluorite)? Will it have a specific property (superconducting or not; ferromagnetic or not)? The tabular data are often plotted as structure maps or phase diagrams, which are easily interpretable in terms of only one (e.g., radius ratio, electron count, or Mendeleev number) or two (e.g., electronegativity and size) parameters.^{20,21} These visual representations are the epitome of simplicity, preferred by practitioners because conclusions can be made quickly by interpreting the patterns on these maps while the complexity of chemical relationships is still captured. For example, the site distributions in normal versus inverse spinels AB_2X_4 can often be explained by crystal-field

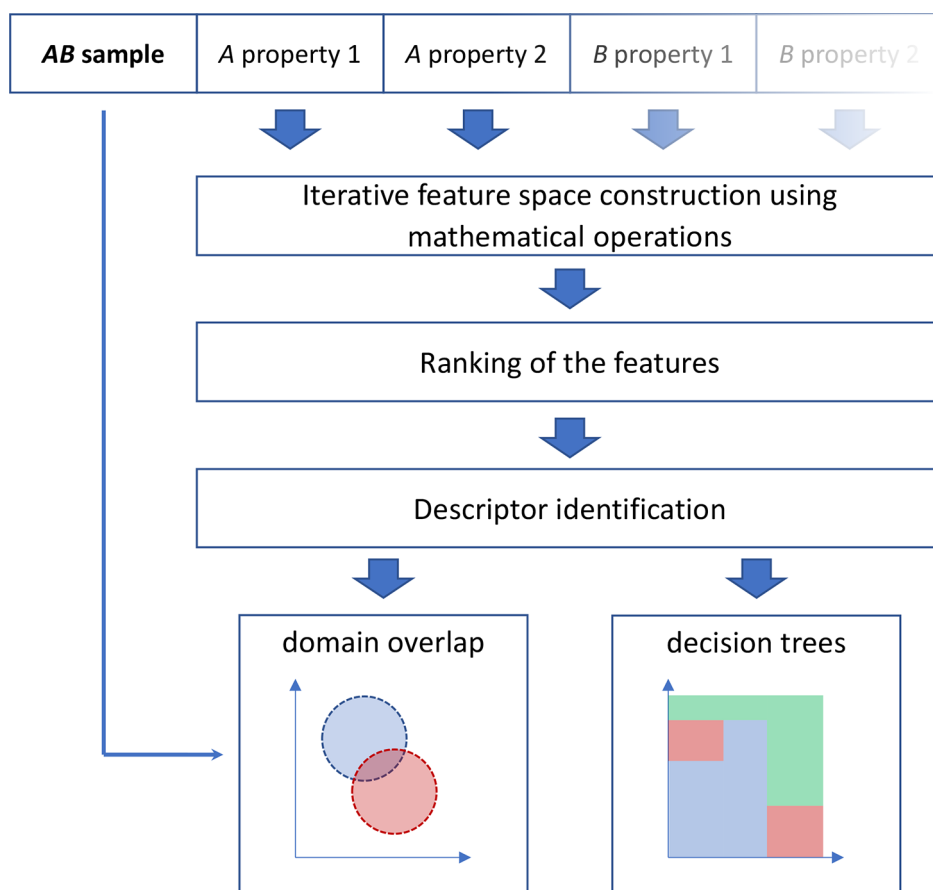


Figure 2. Comparison of the SISSO-DO and SISSO-DT methods.

stabilization energies, but this is only one factor; a structure map based on orbital radii performs the classification more succinctly and accurately (including compounds containing atoms with closed d shells) and reveals that d-orbital interactions are only of minor importance.^{22,23}

Creating effective structure maps requires a good choice of parameters to plot. The recently developed SISSO method, which is based on compressed sensing using symbolic expressions for regression and classification, does just that.²⁴ It builds feature space by combining descriptors with different mathematical operators to generate a large number of equations and selects the best set of equations. A key advantage of this approach is that it is applicable to materials chemistry problems, where the number of samples may be small, while the number of descriptors may be large. In this way, a better tolerance factor, superseding the Goldschmidt tolerance factor, was discovered that predicts the formability of perovskite structures with an accuracy of 92% on a set of 576 ABX₃ compounds.²⁵ This tolerance factor was identified by a SISSO calculation using domain overlap (DO) as the scoring function; the top 100000 descriptors were selected, and a *post hoc* analysis was performed using decision trees (DT) with two terminal leaves. In a new implementation called SISSO++, the final descriptor selection is performed by ranking equations using a support vector machine as well as DO, which can alleviate some of the shortfalls of DO scoring.²⁶ However, as we demonstrate later, useful equations may be discarded during the initial ranking by DO.

We raise the hypothesis that DT may serve as a better scoring function than DO to identify the optimum descriptors.

DT with a shallow depth will be used to screen the millions or billions of descriptors generated recursively in the SISSO feature space. Then, we test this approach on three diverse classification problems in solid-state chemistry (perovskites, spinels, and rare-earth intermetallics) and evaluate how well it performs compared to using DO as the scoring function.

2. METHODS

2.1. Implementation of DT for Scoring Descriptors. In the original SISSO method, the primary features are first combined recursively with mathematical operators to build a large number of equations.²⁴ Then, in a procedure termed “sure independence screening”, the descriptors are scored for classification problems using one-dimensional (1D) DO, in which the domains of samples belonging to each class are approximated by convex hulls and the number of samples that lie in overlapping regions are minimized. In the “sparsifying operator” part of this method, the descriptors are ranked using DO scoring, and a subset of the top *k* features are selected and ordered by complexity to give the best *n*-dimensional descriptors. If several descriptors result in the same degree of convex hull overlap, then the one with minimal *n*-dimensional overlap volume is selected.

In the modified approach proposed here, which pertains only to classification problems, DT was applied to both the sure independence screening and the sparsifying operator parts of the SISSO method. The DT algorithm builds a treelike graph with simple rules to classify samples based on given features; it is nonparametric, simple to interpret, and applicable to multiclass problems. DT was implemented using the Classification and Regression Tree (CART) algorithm.²⁷ The best feature (for descriptors of dimensions greater than one) and threshold splitting points were evaluated using the Gini impurity, $1 - \sum_{i=1}^c P_i^2$, where *c* is the number of classes and *P_i* is the

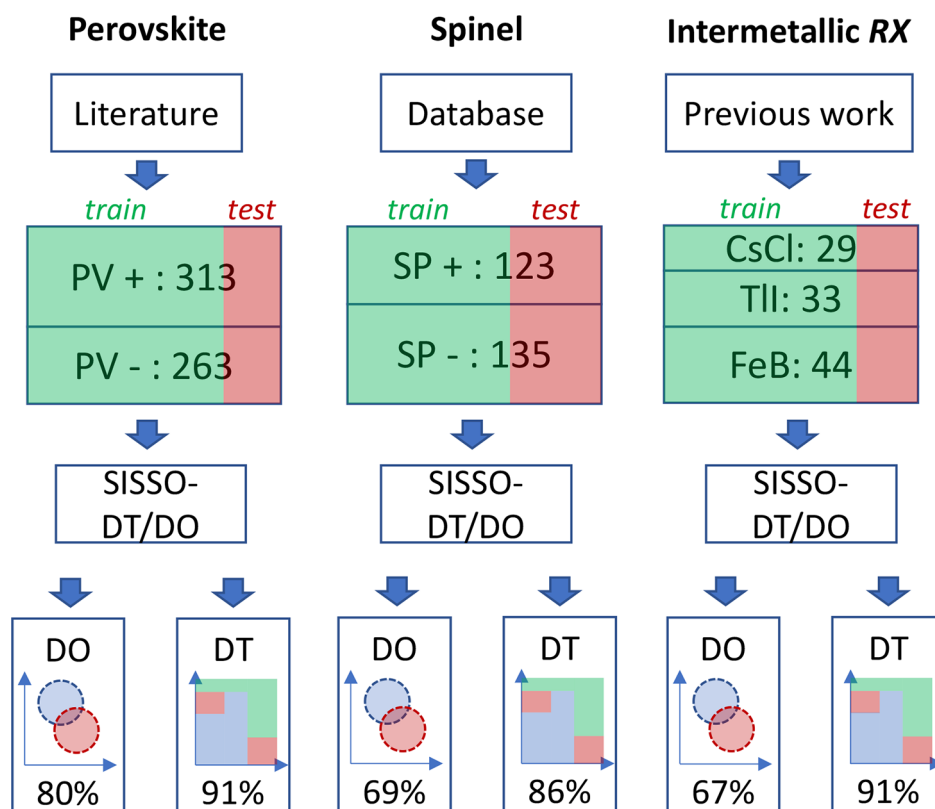


Figure 3. Three data sets examined: ABX_3 compounds classified as perovskites (PV+) or nonperovskites (PV−), AB_2O_4 compounds classified as spinels (SP+) or nonspinel (SP−), and rare-earth intermetallics RX classified as CsCl, TlI, or FeB structures. Models were created using DO or DT as scoring functions to select descriptors by SISSO, and their accuracies in classifying the test sets were compared.

fraction of samples belonging to class i . A node in the DT is split into two child nodes by finding the best feature that minimizes the Gini impurity. The steps are performed recursively until all samples are classified or a stopping criterion is reached. Because shallow trees were used, with a depth of ≤ 4 , no pruning was performed. The accuracy in classifying the training set was used to score the features. This DT algorithm was implemented in Fortran, with the feature ranking and descriptor identification steps modified in the original SISSO code to use trees instead of overlap, as schematized in Figure 2.

2.2. Data Sets. For the illustrative example presented in the Introduction to classify metal oxides MO_2 as rutile or fluorite structures, data were collected from *Pearson's Crystal Data*.¹⁸ Shannon radii and Pauling electronegativity were used as features. In addition to a SISSO model, a random forest model was developed using the Scikit-learn Python library.²⁸

For the main test cases, to compare the two approaches (DO and DT) for descriptor scoring, three data sets were examined. Crystal structure data were gathered for ABX_3 compounds (perovskites vs nonperovskites) from previous literature,²⁵ for oxides AB_2O_4 (spinel vs nonspinel) from *Pearson's Crystal Data*,¹⁸ and for rare-earth intermetallics RX (CsCl, TlI, and FeB structure types) from our own recent work.²⁹ The number of samples in these data sets, a schematic of the workflow, and the classification accuracies for the test sets are summarized in Figure 3.

2.3. Features. Crystal structures depend on many types of features, including the position of an element in the periodic table, which can be described by well-defined quantities (e.g., atomic number Z), or radii and electronegativities, which can be chosen from various scales. To allow a fair comparison, the models were developed with the same features that were most commonly used in previous classification models of perovskites, spinels, and rare-earth intermetallics. For the ABX_3 data set, the features considered were oxidation states (n_A , n_B , and n_X), Shannon ionic radii (r_A , r_B , and r_X),

and radius ratios, with subscripts A, B, and X referring to the component elements. For the AB_2O_4 data set, the features considered were Ahrens radii (r_A^{Ah} , r_B^{Ah} , and r_O^{Ah}), radius ratios, Allred–Rochow electronegativities (χ_A^{AR} , χ_B^{AR} , and χ_O^{AR}), and modified bond stretching force constants (K_{AB}), as defined by:

$$K_{AB} = \frac{\chi_A^{AR} \chi_B^{AR}}{[(r_A^{Ah} + r_B^{Ah})^2 + (r_B^{Ah} + r_X^{Ah})^2 + 1.155(r_A^{Ah} + r_X^{Ah})(r_B^{Ah} + r_X^{Ah})]} \quad (1)$$

Because the atom B in AB_2O_4 can adopt multiple oxidation states in some compounds, indices 1 and 2 were used to identify the lower and higher oxidation states, respectively. For the RX data set, the features considered were atomic numbers of R and X (Z_R and Z_X), metallic radii of R (r_R^m), covalent radii of X (r_X^{co}), Pauling electronegativities of X (χ_X^p), and period and group numbers of X (P_X and G_X). A complete list of these features is provided in Table S1.

Calculations were performed using the original (SISSO-DO, version 3.0.2)²⁴ and modified (SISSO-DT) versions of the code. For SISSO-DO, the top 10000 two-dimensional (2D) descriptors were analyzed *post hoc* using DT. This analysis was performed in Python, using the Decision Tree Classifier implemented in the Scikit-learn package.²⁸ The maximum tree depth was fixed at 3. All computations were performed on a desktop computer equipped with an Intel(R) Core™ i7-10700K and 16 GB of RAM running Ubuntu 20.04. The data sets are provided in the Supporting Information, and the modified SISSO-DT code is available at https://github.com/balaranjan/SISSO_DT.git.

3. RESULTS AND DISCUSSION

3.1. Simulated Data. As a benchmark, a simulated data set was examined using DO versus DT as scoring functions for SISSO analysis. The set consists of 100 randomly generated samples belonging to two classes and distributed over disjoint

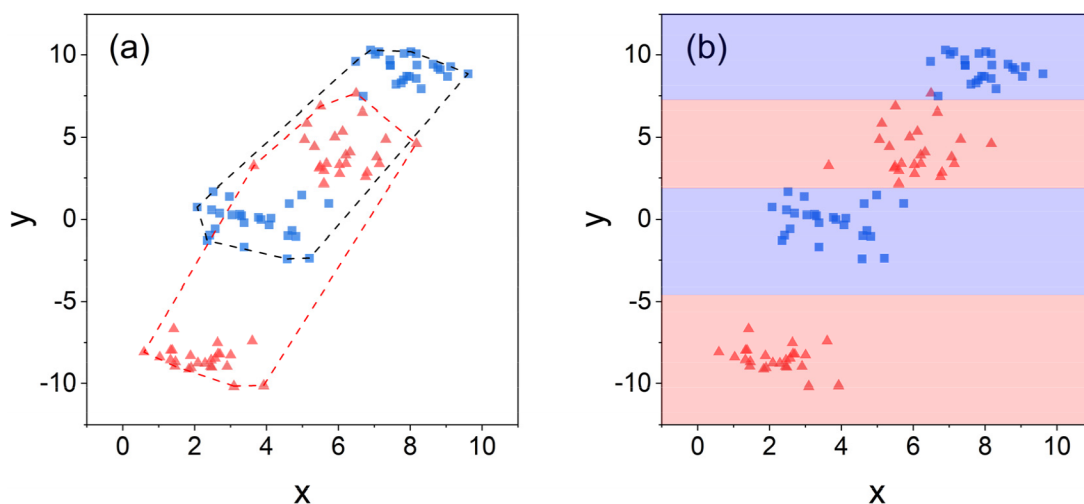


Figure 4. Set of 100 randomly generated samples belonging to two classes (red triangles and blue squares), classified by SISSO using (a) DO and (b) DT. The dashed outlines in part a delineate the convex hull region around each class. The shaded colors in part b indicate the region belonging to each class.

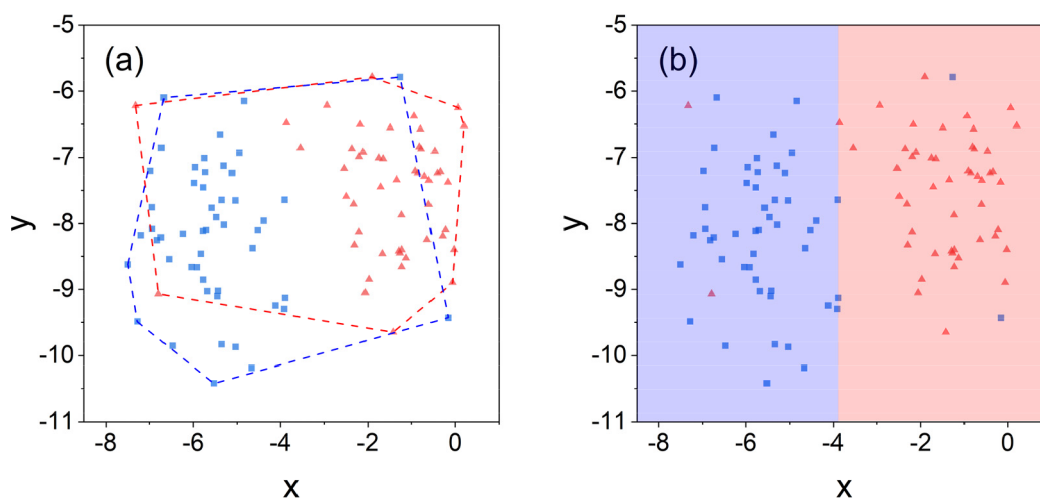


Figure 5. Set of 100 randomly generated samples belonging to two classes and containing four mislabeled samples classified by SISSO using (a) DO and (b) DT.

regions, with their x and y coordinates acting as features. First, SISSO analysis was performed in its original implementation using DO, in which separate convex hulls are created around each class, with the aim to minimize their overlap by optimizing the features. The result is that a cluster of samples belonging to the red class lies within the convex hull of the blue class and vice versa, such that 41% of the samples are misclassified (Figure 4a). Next, SISSO analysis was performed using DT, with a depth of 3. (The depth controls how large a decision tree can grow; larger trees are less interpretable and may lead to overfitting if data sets are small.) Here, different clusters of each class are clearly separated by decision boundaries such that only 1% of the samples are misclassified (Figure 4b).

In general, the success of DO depends strongly on the distribution of samples in the feature space. For well-behaved data, most samples lie within a single cluster close to the center of the convex hull. However, if the samples are distributed over disjoint clusters, the center of the convex hull may not be highly populated, or worse, it could be populated by samples of the opposite class. In such cases, DT may be expected to

perform better, as seen later in rare-earth intermetallics, where outliers in regular trends along a series of f elements are well-known (e.g., europium and ytterbium compounds). As is typical in machine-learning methods, there remains a risk of overfitting, which would be manifested as elaborate convex hull shapes with sharp angles in the DO, or multiple smaller nodes, leading to poor generalizability in DT. A visual assessment of the plotted data goes a long way toward ensuring that the fitting is reasonable.

Next, another simulated data set was examined to evaluate the impact of mislabeled data. Alas, despite best efforts, experimental databases such as *Pearson's Crystal Data*¹⁸ or ICSD³⁰ cannot guarantee 100% trustworthiness. Errors could be introduced at any stage of data collection or processing. They could arise from experimental limitations, related to the instrument or method of analysis, or human foibles, ranging from honest typographical errors to outright malice. Errors can occur with crystallographic assignments (wrong space group, sites, or prototype) or chemical compositions (incorrect or imprecise formula). To illustrate an extreme case, four samples were intentionally mislabeled within the simulated data set. If

DO is used with SISSO, 71% of the samples are misclassified (Figure 5a). As a consequence of these few mislabeled samples (outlying red triangles on the left and blue squares on the right), useful descriptors of lower complexity are unnecessarily discarded, which can lead to overfitting with more complex descriptors. If DT with a depth of 1 is used with SISSO, only these four mislabeled samples are misclassified (Figure 5b).

Errors in a data set introduce a dilemma. A machine-learning practitioner with little background in materials science may inadvertently attribute outliers to imperfections in the model, whereas a materials scientist may be more likely to mistrust the original data and advocate for reinvestigation of the experiments. This scenario demonstrates the need to use scoring functions that are flexible enough to accommodate samples lying in disparate clusters yet belonging to the same class while being robust enough to disregard outliers.

In the following sections, we compare how these two scoring functions for SISSO classification perform when they are applied to three cases of real experimental data. The first two, perovskites and spinels, have been extensively analyzed in the literature previously, whereas the third, rare-earth intermetallics, has not. Two iterations of feature construction (rung = 2) over a subspace of 10000 features were applied. For the DT approach, the maximum tree depth was set to 3 for 2D features. The notation is of the forms O_1^{data} and O_2^{data} for the descriptors derived from overlap scoring and T_1^{data} and T_2^{data} for those derived from DT, where the superscript indicates the data set (PV = perovskites, SP = spinels, and RX = rare-earth intermetallics).

3.2. Perovskites ABX_3 . The data set consists of 576 compounds with the formula ABX_3 , of which 313 are perovskites and 263 are not; there are 369 oxides, 66 fluorides, 71 chlorides, 27 bromides, and 43 iodides.²⁵ It was split into training (80%) and test sets (20%). The top-ranked descriptors selected using DT outperform those using DO scoring (Table 1) and give a more visually striking structure map that places

Table 1. Performance of Models to Classify Perovskites ABX_3 Based on Descriptors Selected by DO and DT Scoring^a

scoring function	descriptor ^b	accuracy	precision	recall
DO	$O_1^{\text{PV}} = \frac{r_A \ln(r_A/r_B)}{n_B}$	0.80 (0.81)	0.79 (0.76)	0.93 (0.93)
	$O_2^{\text{PV}} = \left(r_A \frac{r_A}{r_X} \frac{r_B}{r_X} \right) / n_A$			
DT	$T_1^{\text{PV}} = \frac{r_X}{r_B} - \frac{n_B}{n_X}$	0.91 (0.93)	0.96 (0.95)	0.90 (0.93)
	$T_2^{\text{PV}} = \frac{n_B}{n_X}(r_A - r_B)$			

^aPerformance metrics are shown for test and, in parentheses, training sets. ^bThe descriptors are based on radii (r_A , r_B , and r_X) and oxidation states (n_A , n_B , and n_X).

perovskites within a narrow window (Figure 6). Even with a small number of features (7.9×10^4) involving equations built from two iterations of feature construction, the 2D descriptors (T_1^{PV} and T_2^{PV}) found by DT scoring are able to discriminate well between perovskites and nonperovskites. The accuracy (0.95 for the full data set) is similar to that obtained with the best 2D descriptor found by Bartel et al. (0.95).²⁵ Out of the

576 compounds, 40 were misclassified in the full data set (and 30 in the training set). The descriptor T_1^{PV} does not depend on any feature of element A, which allows the transition points of T_2^{PV} to be tracked individually for different elements B.

These mathematical solutions provide an explanation for classifying perovskites versus nonperovskites, but extracting an interpretation in terms of chemical ideas is another matter wholly. It is appealing when features of each element A, B, and X appear at least once in the descriptors, and when they are related in simple ways such as ratios or differences. The structure map of Figure 6b resembles existing ones built from the Goldschmidt tolerance factor $t = (r_A + r_X)/\sqrt{2}(r_B + r_X)$ and the octahedral factor $\mu = r_B/r_X$, for which there is a narrow stability window for perovskites to form.³¹ It is no accident that the first part of the descriptor T_1^{PV} is r_X/r_B , which is just the inverse of the octahedral factor μ . The second part, n_B/n_X , acts as a correction term that imparts additional restrictions on the relative oxidation states of B and X. Similarly, the descriptor T_2^{PV} expresses a dependence on the relative sizes of A and B but through their difference rather than a ratio, as appears in the Goldschmidt tolerance factor.

3.3. Spinel Oxides AB_2O_4 . A data set of 258 ternary oxides AB_2O_4 was extracted from *Pearson's Crystal Data*, of which 123 adopt the $MgAl_2O_4$ -type structure (hereafter referred to as spinel) and 135 do not.¹⁸ From this data set, 180 entries were used for the training set and 78 for the test set. Scoring with DT results in mathematically simpler descriptors and better performance than that with DO (Table 2). This procedure misclassifies 17 out of 180 entries in the training set and 28 out of 258 in the total data set. The structure map derived from DT shows a much clearer separation between spinels and nonspinel than the one derived from DO (Figure 7).

Many other structure maps to classify spinels have previously been reported in the literature, and it is interesting to compare how well they perform over the same data set. One of the earliest maps, which was developed by Kugimiya and Steinfink, uses the ratio of Ahrens radii $r_A^{\text{Ah}}/r_B^{\text{Ah}}$ and a modified bond stretching constant $K_{AB} \propto \chi_A \chi_B$ as descriptors.³² It misclassifies 60 out of 258 compounds in the total data set. A more recent map by Zhang and Zunger, which uses pseudopotential radii of elements A and B as descriptors, misclassifies 57 compounds in the total data set.²³ Although the more rudimentary models are less accurate, they have the appeal of much simpler descriptors compared to the machine-learning model using DT. In this regard, the machine-learning model is somewhat disadvantaged in being less easily interpretable even if it uses similar elemental features.

At first glance, the descriptors derived from SISSO-DT may appear opaque, but a more careful examination shows that they incorporate structurally and chemically meaningful information consistent with previously developed structure maps. The first descriptor T_1^{SP} contains expressions relating electronegativities in the numerator ($\chi_A^{\text{AR}} + \chi_O^{\text{AR}}$) and radii in the denominator ($r_A^{\text{Ah}}/r_B^{\text{Ah}}$), both of which are captured in the bond stretching constant K_{AB} in similar ways and account for electronic and structural effects. The second descriptor T_2^{SP} contains the radius ratio $r_A^{\text{Ah}}/r_B^{\text{Ah}}$, which is equivalent to the index used in the structure map of Kugimiya and Steinfink.³² This ratio is also (inversely) proportional to a geometric tolerance factor recently proposed by Song and Liu for normal spinels, $\tau = \sqrt{3}(r_B + r_X)/2(r_A + r_X)$.³³ By itself, this tolerance factor is insufficiently versatile to distinguish spinels from other

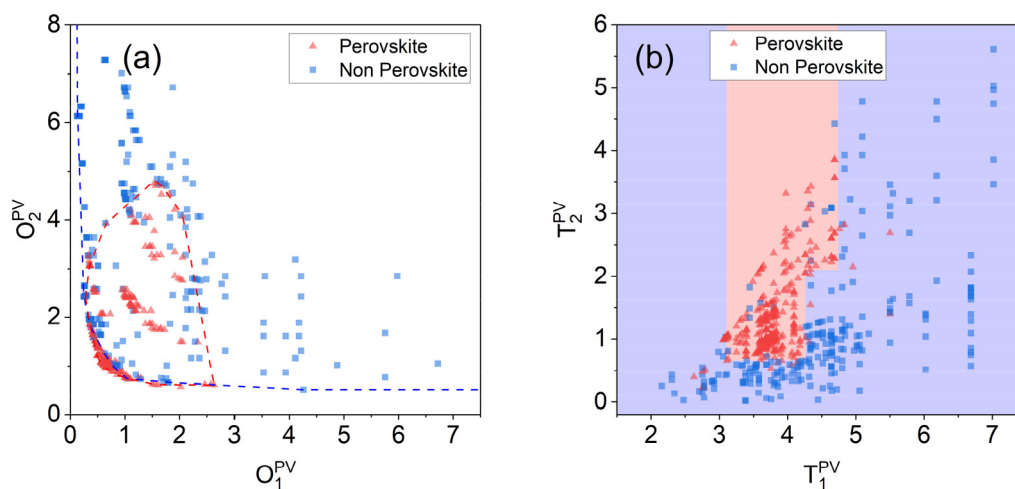


Figure 6. Classification of perovskites (red triangles) and nonperovskites (blue squares) among ABX_3 compounds using the best descriptors selected by (a) DO and (b) DT as scoring functions.

Table 2. Performance of Models to Classify Spinel Oxides AB_2O_4 Based on Descriptors Selected by DO and DT Scoring^a

scoring function	descriptor ^b	accuracy	precision	recall
DO	$O_1^{SP} = \frac{\exp(r_{B2}^{Ah})r_{B2}^{Ah}\chi_A^{AR}}{r_A^{Ah}}$	0.69 (0.74)	0.63 (0.66)	0.89 (0.91)
	$O_1^{SP} = \frac{\frac{r_A^{Ah}}{r_{B1}^{Ah}} + \frac{r_{B1}^{Ah}}{r_O^{Ah}}}{\frac{r_{B2}^{Ah}\chi_A^{AR}}{r_O^{Ah}}}$			
DT	$T_1^{SP} = \frac{r_O^{Ah}(\chi_A^{AR} + \chi_O^{AR})}{r_A^{Ah}r_{B1}^{Ah}}$	0.86 (0.91)	0.85 (0.87)	0.87 (0.94)
	$T_2^{SP} = \frac{K_{AB2}}{r_A^{Ah}} \left(\frac{r_A^{Ah}}{r_{B2}^{Ah}} \right)^3$			

^aPerformance metrics are shown for test and, in parentheses, training sets. ^bThe descriptors are based on Ahrens radii (r_A , r_B , and r_O), Allred–Rochow electronegativities (χ_A , χ_B , and χ_O), and a bond stretching constant K_{AB} .

AB_2O_4 structure types, as can be seen by the significant overlap (Figure 8a). Nevertheless, the appeal of a 1D descriptor to classify spinels is undeniable.

The insight gained above suggests that the key expressions related to radii ($r_A^{Ah}r_{B1}^{Ah}$ and r_A^{Ah}/r_{B2}^{Ah}) can be put to advantage. In a simpler structure map created using these descriptors, spinels are located in a cluster centered at $r_A^{Ah}r_{B1}^{Ah} = 0.45$ and $r_A^{Ah}/r_{B2}^{Ah} = 1.1$ (Figure 8b). If normal spinel structures are assumed, Pauling rules indicate that $r_A^{Ah}r_{B1}^{Ah}$ values are expected to lie between 0.18 and 0.48. For most spinels, the value of r_A^{Ah}/r_{B2}^{Ah} is greater than 0.8, indicating a preference for larger A cations relative to B cations. Most spinels fall within the ellipse, allowing a 1D tolerance factor to be defined as

$$\tau_s = 2.5(r_A^{Ah}r_{B2}^{Ah} - 0.45)^2 + 0.3\left(\frac{r_A^{Ah}}{r_{B1}^{Ah}} - 1.2\right)^2 - 0.1 \quad (2)$$

with the condition $\tau_s \leq 0.0$ and an accuracy of 84% in classification.

Questions may be raised about how sensitive the machine-learning models are to different types of features and how they would perform on imbalanced data sets. If Ahrens radii are

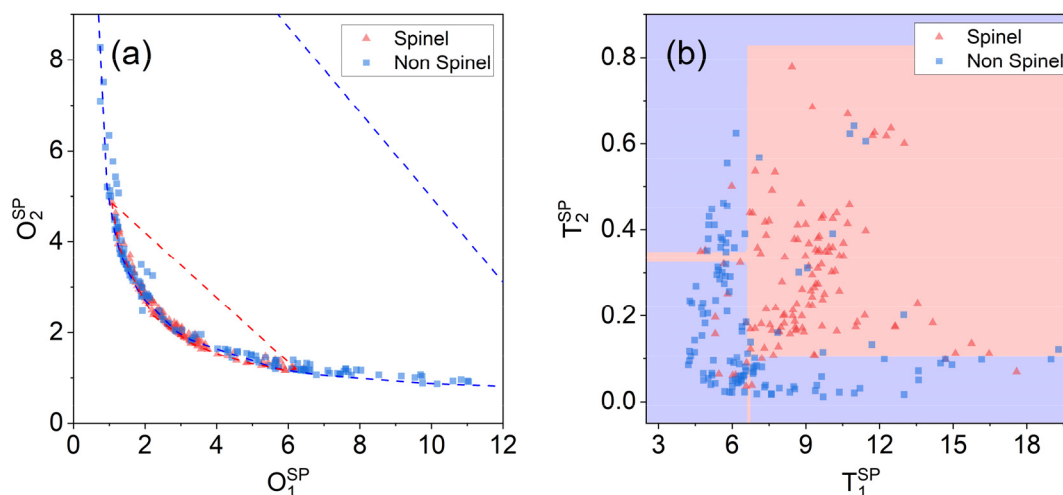


Figure 7. Classification of spinels (red triangles) and nonspinel (blue squares) among AB_2O_4 compounds using the best descriptors selected by (a) DO and (b) DT as scoring functions.

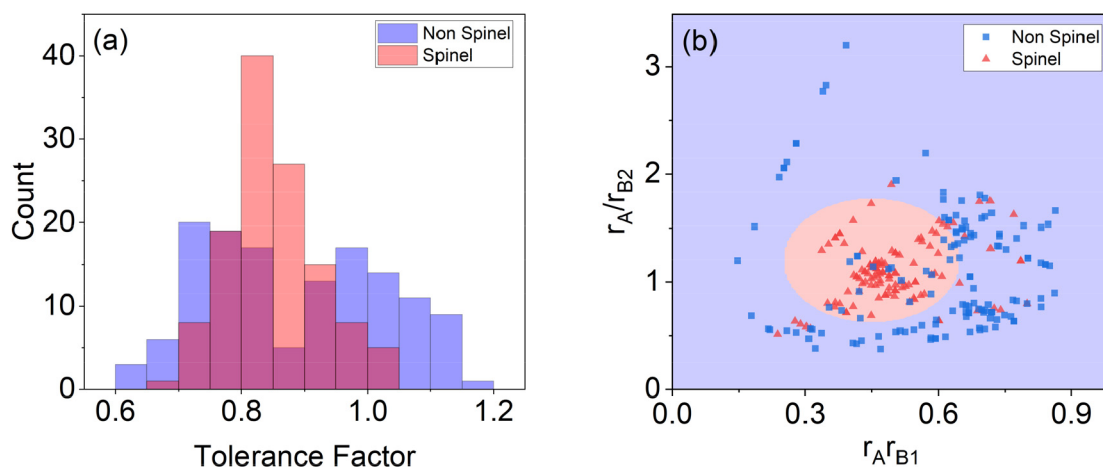


Figure 8. Classification of spinels by (a) the tolerance factor from Song and Liu and (b) the Ahrens radii of A and B cations (spinels as red triangles; nonspinels as blue squares). The shaded ellipse in part b classifies spinels with 84% accuracy.

replaced by Shannon radii in the descriptors listed in Table 2, DT scoring still performs better than DO scoring (Table S2). The oxides AB_2O_4 form a balanced data set for which machine-learning methods are expected to behave well. In contrast, the chalcogenides AB_2X_4 ($X = S, Se, Te$) form an imbalanced data set containing half as many spinels as nonspinels, as well as many more sulfides and selenides than tellurides. When the two methods are applied to this imbalanced data set, DT scoring still achieves a higher accuracy (0.74) than DO scoring (0.59) (Table S3).

3.4. Rare-Earth Intermetallics RX. The data set consists of 106 equiatomic rare-earth intermetallics RX ($X = Si, Au, Cu, Ag, Rh, Pd, Ni, Pt$) extracted from *Pearson's Crystal Data*.¹⁸ These compounds were examined because the distribution of the three structure types adopted (44 FeB, 33 TiI, and 29 CsCl) appears to be puzzlingly haphazard. Similar to how structure maps for other classes of compounds were historically developed through a combination of trial-and-error and chemical intuition, the elements R and X were rearranged in different sequences until a reasonable separation of these structure types was achieved.²⁹ A good descriptor for the x axis of the RX structure map was found to be a simple combination of R-element features, $x = Zr^m$, where Z is the atomic number and r^m is the metallic radius. However, the functional form of the descriptor for the y axis was not obvious. To determine this descriptor, numerical values were assigned to the X element ($Si = 1, \dots, Pt = 8$) and SISSO regression was applied, giving

$$y = \frac{\sqrt{Z} \times P/G}{\exp(\chi_p r_{co})} \quad (3)$$

where Z is the atomic number, P is the period, G is the group, χ_p is the Pauling electronegativity, and r_{co} is the covalent radius of the X element. The final mathematical expression obtained in this process depends on the depth and complexity parameters set in the algorithm, which requires some judgment to be exercised by the user. This is why a visual inspection of the structure map by a domain expert, a materials scientist, in this case, can be valuable. Only 2 out of 106 entries in the data set are misclassified in this structure map, and its success in explaining many experimental results, such as polymorphism and stabilization of a different structure by adding a third component, lends credence to it. It is noteworthy that SISSO classification with scoring by DO fails to give a satisfactory

structure map when the same set of starting features was used and even when the good descriptors obtained above were fed in. As discussed earlier, DO scoring tends to perform poorly when samples belonging to the same class are distributed over disjointed clusters in feature space, as is the case here. Thus, it is interesting to determine whether DT scoring will do a better job.

The data set of these 106 RX intermetallics was split into training (80%) and test (20%) sets, with stratified splitting applied to maintain equitable class distributions of the three structure types (FeB, TiI, and CsCl) within these sets. A tree depth of 2 was set for DT scoring of the feature space. (When three classes are present, more than two leaf nodes in the DT are needed for classification.) With the descriptors obtained by DT scoring, the performance is significantly better than that by DO scoring (Table 3). The recall is also higher here, in

Table 3. Performance of Models to Classify Rare-Earth Intermetallics RX Based on Descriptors Selected by DO and DT Scoring^a

scoring function	descriptor ^b	accuracy	precision	recall
DO	$O_1^{RX} = \frac{PXr_R^m}{\ln(r_X^{co})}$	0.67 (0.84)	0.65 (0.77)	0.62 (0.86)
	$O_2^{RX} = \frac{G_X r_R^m}{Z_R}$			
DT	$T_1^{RX} = \exp(\chi_X^p) G_X r_R^m$	0.91 (0.95)	0.92 (0.95)	0.91 (0.95)
	$T_2^{RX} = \ln(Z_X) Z_R r_X^{co}$			

^aPerformance metrics are shown for test and, in parentheses, training sets. ^bThe descriptors are based on atomic number Z , period P , group G , Pauling electronegativity χ_p , and covalent radius r_{co} .

contrast to the previous two case studies, in which this metric was slightly inferior using DT scoring. The explanation can be traced to the significant overlap of the convex hull for FeB with those of the other two structure types, with samples clearly seen to be distributed over disjointed regions (Figure 9a), which poses no problem for the structure map based on DT scoring (Figure 9b). The DT model misclassifies 4 entries in the training set and 2 in the test set.

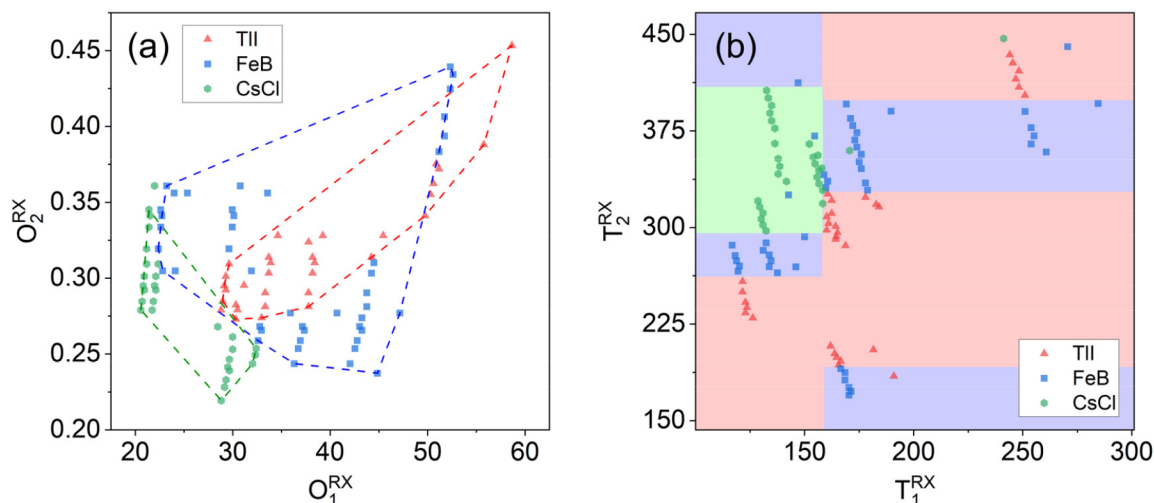


Figure 9. Classification of TII (red triangles), FeB (blue squares), and CsCl (green circles) structure types adopted by rare-earth intermetallics RX using the best descriptors selected by (a) DO and (b) DT as scoring functions.

Table 4. *Post Hoc* Analysis of SISSO-DO Descriptors Selected by DT^a

data set	descriptor	accuracy	precision	recall
perovskites ABX ₃	$O_3^{PV} = \frac{r_A r_X}{(r_B)^2} \left/ \left(\frac{r_A}{r_B} - \frac{r_B}{r_X} \right) \right.$ $O_4^{PV} = \frac{n_A n_B}{\exp(r_B)}$	0.85 (0.92)	0.91 (0.94)	0.85 (0.90)
spinel oxides AB ₂ O ₄	$O_3^{SP} = \frac{(\chi_B^{AR})^2 r_O^{Ah} r_A^{Ah}}{(r_{B2}^{Ah})^2}$ $O_4^{SP} = \frac{(r_B^{Ah})^2}{r_O^{Ah} r_A^{Ah}} + \frac{\chi_O^{AR}}{\chi_A^{AR}}$	0.69 (0.87)	0.71 (0.91)	0.63 (0.81)
rare-earth intermetallics RX	$O_3^{RX} = (G_X \chi_X^P)(r_R^m - r_X^{co})$ $O_4^{RX} = (Z_R + Z_X) \frac{G_X}{P_X}$	0.68 (0.90)	0.69 (0.90)	0.68 (0.90)

^aPerformance metrics are shown for test and, in parentheses, training sets.

3.5. Post Hoc Analysis. The results demonstrate that DT serves as a scoring function better than that of DO to classify perovskites, spinel oxides, and rare-earth intermetallics. For the SISSO-DO method, the performance metrics (accuracy, precision, recall) were calculated using DO. As a fair comparison, it would be interesting to evaluate the performance of the descriptors selected by SISSO-DO using DT. To this end, the top 10,000 2D descriptors from the SISSO-DO models were ranked by DT. The maximum tree depth was fixed at 3 and the descriptor pairs in SISSO-DO giving highest accuracy on the training sets were examined (Table 4).

For the perovskite data set, the DT-selected descriptor pair (O_3^{PV} , O_4^{PV}) performs significantly better than the SISSO-DO descriptor pair (O_1^{PV} , O_2^{PV}) (Table 1), with an increase of 0.05 in accuracy in the test set. However, this pair still falls short compared to the SISSO-DT pair by 0.06 in test set accuracy. For the spinel oxides and rare-earth intermetallics data sets, there is no significant difference (0.01 or less) in test set accuracy between the DT-selected and SISSO-DO descriptor pairs (Tables 2 and 3), but they are far outperformed by the SISSO-DT pairs. Overall, the SISSO-DT descriptors give accuracies of ≥ 0.91 on the training sets and ≥ 0.86 on the test sets for the three classification problems, while using a relatively smaller feature space. Shallow DT (having low

depths) does not hamper interpretability. In addition to improvements in accuracy, SISSO-DT is generally faster than SISSO-DO. For the perovskite data set, which has the greatest number of samples, SISSO-DT required 140 min, whereas SISSO-DO required 262 min. For the spinel data set, which has the greatest number of features, SISSO-DT required 40 min and SISSO-DO required 51 min. Finally, although it might be possible to increase the dimensions of the descriptor space to improve the accuracy of the machine-learning models, this will come at the expense of decreased interpretability and significantly higher computational costs.

4. CONCLUSIONS

DT was proposed, implemented, and tested to score features generated by the SISSO method. To classify crystal structures of perovskites, spinel oxides, and rare-earth intermetallics, SISSO-DT offers significantly improved performance using features that are better than those of SISSO-DO. The visualization of these machine-learning models, as manifested in structure maps based on chemical descriptors related in simple ways, facilitates interpretation and enables experimentalists to make predictions of new materials with greater confidence.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.inorgchem.3c01153>.

Description of features, models to classify spinel oxides using Shannon radii, and models to classify spinel chalcogenides (PDF)

Data set for perovskites (Perovskites_pred.xlsx) (XLSX)

Data set for spinel oxides (Spinel_Oxide_pred.xlsx) (XLSX)

Data set for binary rare-earth intermetallics (RX_pred.xlsx) (XLSX)

Data set for spinel chalcogenides (Spinel_Chalk_pred.xlsx) (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Arthur Mar – Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada; orcid.org/0000-0003-0474-5918; Email: amar@ualberta.ca

Authors

Balaranjan Selvaratnam – Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada
Anton O. Oliynyk – Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada; Department of Chemistry and Biochemistry, Manhattan College, Riverdale, New York 10471, United States; orcid.org/0000-0003-0732-7340

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.inorgchem.3c01153>

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (through Discovery Grant RGPIN-2018-04294), the Canada First Research Excellence Fund (through the Future Energy Systems Research Institute at the University of Alberta under Project T12-P01), and resources available through Digital Research Alliance of Canada (alliancecan.ca).

■ REFERENCES

- (1) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559*, 547–555.
- (2) Lapuschkin, S.; Wäldchen, S.; Binder, A.; Montavon, G.; Samek, W.; Müller, K.-R. Unmasking Clever Hans predictors and assessing what machines really learn. *Nat. Commun.* **2019**, *10*, 1096.
- (3) Saal, J. E.; Oliynyk, A. O.; Meredig, B. Machine learning in materials discovery: Confirmed predictions and their underlying approaches. *Annu. Rev. Mater. Res.* **2020**, *50*, 49–69.
- (4) Wang, A. Y.-T.; Murdock, R. J.; Kauwe, S. K.; Oliynyk, A. O.; Gurlo, A.; Brgoch, J.; Persson, K. A.; Sparks, T. D. Machine learning for materials scientists: An introductory guide toward best practices. *Chem. Mater.* **2020**, *32*, 4954–4965.
- (5) Horton, M. K.; Dwaraknath, S.; Persson, K. A. Promises and perils of computational materials databases. *Nat. Comput. Sci.* **2021**, *1*, 3–5.
- (6) Chuang, K. V.; Keiser, M. J. Adversarial controls for scientific machine learning. *ACS Chem. Biol.* **2018**, *13*, 2819–2821.
- (7) Ribeiro, M. T.; Singh, S.; Guestrin, C. “Why should I trust you?” Explaining the predictions of any classifier. *arXiv preprint* **2016**.
- (8) Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215.
- (9) Oviedo, F.; Ferres, J. L.; Buonassisi, T.; Butler, K. T. Interpretable and explainable machine learning for materials science and chemistry. *Acc. Mater. Res.* **2022**, *3*, 597–607.
- (10) Rudin, C.; Chen, C.; Chen, Z.; Huang, H.; Semenova, L.; Zhong, C. Interpretable machine learning: Fundamental principles and 10 grand challenges. *Statist. Surv.* **2022**, *16*, 1–85.
- (11) Esterhuizen, J. A.; Goldsmith, B. R.; Linic, S. Interpretable machine learning for knowledge generation in heterogeneous catalysis. *Nat. Catal.* **2022**, *5*, 175–184.
- (12) Fisher, A.; Rudin, C.; Dominici, F. All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *J. Mach. Learn. Res.* **2019**, *20*, 1–81.
- (13) Lundberg, S. M.; Lee, S.-I. A unified approach to interpreting model predictions. *Proc. 31st Conf. Neural Information Process. Syst.* **2017**, 4768–4777.
- (14) Omidvar, N.; Pillai, H. S.; Wang, S.-H.; Mou, T.; Wang, S.; Athawale, A.; Achenie, L. E. K.; Xin, H. Interpretable machine learning of chemical bonding at solid surfaces. *J. Phys. Chem. Lett.* **2021**, *12*, 11476–11487.
- (15) George, J.; Hautier, G. Chemist versus machine: Traditional knowledge versus machine learning techniques. *Trends Chem.* **2021**, *3*, 86–95.
- (16) Adams, D. M. *Inorganic Solids: An Introduction to Concepts in Solid-state Structural Chemistry*; Wiley: London, 1974; pp 99–100.
- (17) Wells, A. F. *Structural Inorganic Chemistry*, 5th ed.; Clarendon Press: Oxford, U.K., 1984; pp 317–318.
- (18) Villars, P.; Cenzual, K. *Pearson's Crystal Data—Crystal Structure Database for Inorganic Compounds (on DVD)*, release 2021/22; ASM International: Materials Park, OH, 2021.
- (19) Pauling, L. The sizes of ions and the structure of ionic crystals. *J. Am. Chem. Soc.* **1927**, *49*, 765–790.
- (20) Pettifor, D. G. The structures of binary compounds: I. Phenomenological structure maps. *J. Phys. C: Solid State Phys.* **1986**, *19*, 285–313.
- (21) Villars, P. A three-dimensional structural stability diagram for 998 binary AB intermetallic compounds. *J. Less-Common Met.* **1983**, *92*, 215–238.
- (22) Burdett, J. K.; Price, G. D.; Price, S. L. Role of the crystal-field theory in determining the structures of spinels. *J. Am. Chem. Soc.* **1982**, *104*, 92–95.
- (23) Zhang, X.; Zunger, A. Diagrammatic separation of different crystal structures of A_2BX_4 compounds without energy minimization: A pseudopotential orbital radii approach. *Adv. Funct. Mater.* **2010**, *20*, 1944–1952.
- (24) Ouyang, R.; Curtarolo, S.; Ahmetcik, E.; Scheffler, M.; Ghiringhelli, L. M. SISO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Phys. Rev. Mater.* **2018**, *2*, 083802.
- (25) Bartel, C. J.; Sutton, C.; Goldsmith, B. R.; Ouyang, R.; Musgrave, C. B.; Ghiringhelli, L. M.; Scheffler, M. New tolerance factor to predict the stability of perovskite oxides and halides. *Sci. Adv.* **2019**, *5*, No. eaav0693.
- (26) Purcell, T. A. R.; Scheffler, M.; Carbogno, C.; Ghiringhelli, L. M. SISO++: A C++ implementation of the sure-independence screening and sparsifying operator approach. *J. Open Source Softw.* **2022**, *7*, 3960.
- (27) Breiman, L. *Classification and Regression Trees*; Routledge: New York, 2017.

- (28) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- (29) Gvozdetzkyi, V.; Selvaratnam, B.; Oliynyk, A. O.; Mar, A. Revealing hidden patterns through chemical intuition and interpretable machine learning: A case study of binary rare-earth intermetallics *RX*. *Chem. Mater.* **2023**, *35*, 879–890.
- (30) Bergerhoff, G.; Brown, I. D. Inorganic crystal structure database. In *Crystallographic Databases*; Allen, F. H., Bergerhoff, G., Sievers, R., Eds.; International Union of Crystallography: Chester, U.K., 1987.
- (31) Tilley, R. J. D. *Perovskites: Structure-Property Relationships*; Wiley: Chichester, U.K., 2016.
- (32) Kugimiya, K.; Steinfink, H. The influence of crystal radii and electronegativities on the crystallization of AB_2X_4 stoichiometries. *Inorg. Chem.* **1968**, *7*, 1762–1770.
- (33) Song, Z.; Liu, Q. Tolerance factor and phase stability of the normal spinel structure. *Cryst. Growth Des.* **2020**, *20*, 2014–2018.