The Journal of Chemical Physics

A rapid feature selection method for catalyst design: Iterative Bayesian additive regression trees (iBART)

Cite as: J. Chem. Phys. **156**, 164105 (2022); https://doi.org/10.1063/5.0090055
Submitted: 02 March 2022 • Accepted: 04 April 2022 • Accepted Manuscript Online: 08 April 2022 • Published Online: 27 April 2022

Chun-Yen Liu, Shengbin Ye, Meng Li, et al.

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

A rapid feature selection method for catalyst design: Iterative Bayesian additive regression trees (iBART)

Cite as: J. Chem. Phys. 156, 164105 (2022); doi: 10.1063/5.0090055
Submitted: 2 March 2022 • Accepted: 4 April 2022 •
Published Online: 27 April 2022

Chun-Yen Liu,[1] Shengbin Ye,[2] Meng Li,[2,a)] and Thomas P. Senftle[1,a)]

AFFILIATIONS

[1] Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, USA
[2] Department of Statistics, Rice University, Houston, Texas 77005, USA

Note: This paper is part of the JCP Special Topic on Chemical Design by Artificial Intelligence.
[a)] **Authors to whom correspondence should be addressed:** meng@rice.edu and tsenftle@rice.edu

ABSTRACT

Feature selection (FS) methods often are used to develop data-driven descriptors (i.e., features) for rapidly predicting the functional properties of a physical or chemical system based on its composition and structure. FS algorithms identify descriptors from a candidate pool (i.e., feature space) built by feature engineering (FE) steps that construct complex features from the system's fundamental physical properties. Recursive FE, which involves repeated FE operations on the feature space, is necessary to build features with sufficient complexity to capture the physical behavior of a system. However, this approach creates a highly correlated feature space that contains millions or billions of candidate features. Such feature spaces are computationally demanding to process using traditional FS approaches that often struggle with strong collinearity. Herein, we address this shortcoming by developing a new method that interleaves the FE and FS steps to progressively build and select powerful descriptors with reduced computational demand. We call this method iterative Bayesian additive regression trees (iBART), as it iterates between FE with unary/binary operators and FS with Bayesian additive regression trees (BART). The capabilities of iBART are illustrated by extracting descriptors for predicting metal–support interactions in catalysis, which we compare to those predicted in our previous work using other state-of-the-art FS methods (i.e., least absolute shrinkage and selection operator + l_0, sure independence screening and sparsifying operator, and Bayesian FS). iBART matches the performance of these methods yet uses a fraction of the computational resources because it generates a maximum feature space of size $O(10^2)$, as opposed to $O(10^6)$ generated by one-shot FE/FS methods.

I. INTRODUCTION

Statistical learning (SL) and machine learning (ML) have risen in the chemistry and materials science communities as popular tools to for constructing predictive models from physical descriptors.[1–5] Application examples are far-ranging, such as studies that use ML to navigate complex reaction networks[6,7] or to design materials by high throughput screening.[8–11] Beyond prediction, another powerful aspect of SL and ML is the potential for deriving new scientific knowledge by building interpretable models from complex datasets. Our group has focused on applying feature selection (FS) techniques to better understand and predict metal–support interactions in catalysis. We demonstrated that SL methods can derive robust descriptors to quantify the binding energy of single metal atoms on various oxide supports, where the SL descriptors were constructed from primary features that quantify charge transfer between the supported metal atom and the oxide surface.[12,13] In this work, we develop an improved FS technique, called iterative Bayesian additive regression trees (iBART), that is more robust and less computationally demanding than the methods employed in our previous studies.

FS techniques differ from nonparametric methods, such as decision tree methods,[14–16] Gaussian process (GP) methods,[10,17–19] and neural networks,[20–22] which construct predictive models that are powerful but are difficult to interpret.[12,13,23–28] In general, FS approaches begin with the collection of fundamental properties that the user—based on physical intuition—believes are relevant to the property of interest.[23] In the single-metal-atom catalyst example,[12] we collected properties of the supported metals (e.g., ionization energy and electron affinity) and the oxide surfaces (e.g.,

work function and oxygen vacancy formation energy) to build a pool of primary features that would be engineered to construct a feature space of complex descriptors. These properties were chosen because they quantify the ability of the metal or surface to donate or accept charge, which, in turn, influences the metal–oxide binding strength.[29] Although related, the binding energy does not correlate linearly with these simple primary features, and, thus, more complex features are needed to build a physically accurate model. As such, the primary features serve as the basis for feature engineering (FE), whereby a series of unary and binary mathematical operators (i.e., square, inverse, and multiplication) act on the primary features to produce a secondary feature space containing complex functional forms involving non-linear combinations of the primary features. The FE process can be repeated an arbitrary number of times to generate tertiary features, quaternary features, and so on until the feature space contains features with sufficient complexity to predict the property of interest in a linear model. The resulting feature space may include millions or billions of candidate features, and it is the task of the FS algorithm, such as principle component analysis (PCA),[30] recursive feature elimination (RFE),[31] least absolute shrinkage and selection operator (LASSO),[12,23,28] sure independence screening and sparsifying operator (SISSO),[25,27,32] or Bayesian FS methods (e.g., Horseshoe prior and Dirichlet–Laplace prior[13]), to identify the handful of features that best predict the property of interest. In this way, the sequential one-shot FE → FS approach generates models that relate basic properties of a system's individual components to complex properties of the aggregate system in an interpretable functional form.

The one-shot FE → FS approach is generally robust and has been demonstrated in many studies, such as those building models that predict whether a material composition will form a zinc blende or a rock salt structure,[23] predict adsorbate binding energies on various morphologies of bimetallic metal alloys,[27] or connect single-atom metal–support binding energies to metal sintering rates.[28] However, there are several drawbacks and limitations inherent to the one-shot FE → FS method. Recursive FE is necessary to generate features with sufficient complexity but inevitably leads to a high number of candidate descriptors (i.e., $p \sim 10^6$). In chemistry and materials science, the number of candidate descriptors typically is orders of magnitude higher than the size of the available training dataset (i.e., $n \sim 10^2$). Furthermore, candidate descriptors are created from the same basis of primary features and, as such, they are highly correlated with each other. The one-shot FE approach creates a significant challenge for the FS procedure, which is tasked with identifying descriptors from a large pool of correlated candidates with limited training data (i.e., the left path in Fig. 1). As such, one-shot approaches suffer from a high rate of false positives, and significant computational resources are required to handle the enormous feature space they generate. In addition, available memory becomes a bottleneck that limits the number of FE steps that can be applied, thus limiting the complexity of the candidate features.

We address these problems by developing an iterative FE → FS → FE → FS → ⋯ approach, which we call iBART. The mathematical development of iBART is reported in another publication;[33] here, we will focus on the application of iBART in the context of understanding metal–support interactions in catalysis. iBART eliminates the bottleneck created by the one-shot

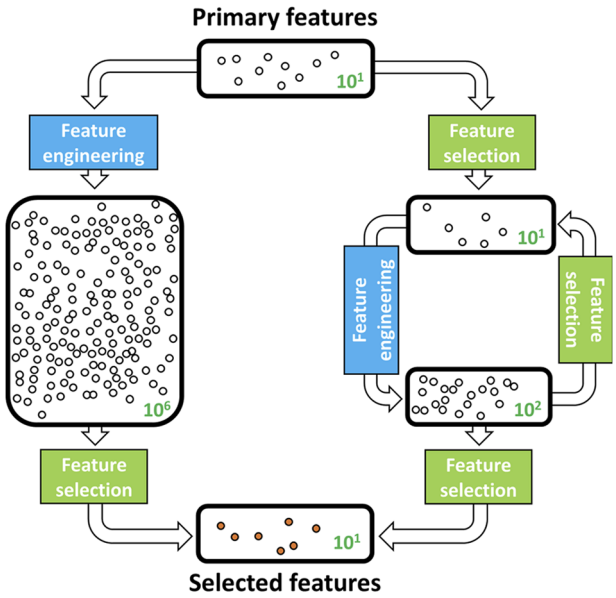

FIG. 1. Illustration of a one-shot FE → FS strategy (left path) and an iterative FE → FS → FE → FS →. . . strategy (right path). The boxes represent the feature space, and the number of spheres inside the boxes indicates the quantity of features in the space.

FE → FS approach by interleaving multiple FE steps with the FS steps (i.e., the right-hand path in Fig. 1). This allows the method to curate the feature space as it is built, which limits the maximum size of the feature space to $O(10^2)$–$O(10^3)$ candidates, instead of $\sim O(10^6)$ generated by the one-shot methods. iBART maintains high representation power and selection accuracy by utilizing flexible statistical FS modules to select features that are nonlinearly correlated with the property of interest, followed by linear FS to form an interpretable model. In particular, the FS steps are executed by a combination of Bayesian additive regression trees (BART)[34,35] and LASSO[36] regression, as described fully in our separate publication.[33]

In this work, we test iBART by applying it to derive models based on our existing datasets describing metal–support interactions in catalysis, which we then compare to models derived using existing FS methods (i.e., LASSO + l_0,[12,23] SISSO,[25] and Dirichlet–Laplace prior[13]). Metal–support interactions, which we quantify here as the binding strength between supported metal atoms and oxide surfaces, are an important metric for designing sinter-resistant single-atom catalysts.[12,29,37,38] If we can estimate the binding energy of the metal atom, the metal cluster size distributions can then be predicted with kinetic sintering models that derive cluster morphology as a function of temperature.[28] We show that iBART matches the accuracy of the other methods while requiring a fraction of the computational resources. In addition, we find that iBART can derive a rich menu of functional forms in the final physical models, which can be a useful aid when interpreting the underlying physics described by the models.

II. METHODS

The objective of iBART in this work is to derive a model for predicting metal/oxide binding energies by using FE and FS to construct a predictive equation that takes the following form:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_D x_D, \quad (1)$$

where $\hat{y}$ is the predicted metal binding energy, $x_1, x_2, \ldots, x_D$ are the selected features composed of properties of the system's components [e.g., electron affinity of the adsorbed metal (EA^m) and oxygen vacancy formation energy of the oxide support (O_{vac})], β_0 is a constant intercept, $\beta_1, \ldots, \beta_D$ are weighting coefficients for each selected feature, and D is the number of features used in the predictive model (i.e., the dimension of the model). The choice of primary features, mathematical operators in the FE step, algorithm in the FS step, and evaluation methods for assessing model performance are defined in Secs. II A–II D.

A. Datasets

In this work, we focus primarily on evaluating the ability of iBART to predict physical descriptors in the context of catalyst design. As such, we use iBART to construct predictive models on two real datasets containing metal–support binding energies from our previously published work, which here we refer to as the metal/support (M–S) dataset[12] and the metal/doped-support (M–D) dataset.[13] The dependent variable [i.e., $\hat{y}$ in Eq. (1)] is the binding energy of single metal atoms (e.g., Au, Cu, and Ir) on various oxide supports [e.g., MgO(100), CeO_2(111), and TiO_2(011)] for the M–S dataset and is the change in binding energy of single metal atoms on MgO(100) surfaces after they are modified with a dopant [e.g., Al-doped or Na-doped MgO(100) surfaces] for the M–D dataset. The primary features are electronic properties of the system's components (i.e., adsorbed metal, oxide support, and surface dopant), such as ionization energy, atomic radius, and atomic number. The datasets and their complete descriptions are available in our previous work.[12,13]

B. Feature engineering

FE is applied to introduce non-linear correlations between the primary features and the property of interest. Unary and binary FE steps are applied to introduce non-linear dependencies and complexity in the candidate features. The set of unary FE operators is $O_1 = \{I, ^{-1}, ^2, \sqrt{}, \log, \exp, |\cdot|, \sin(\pi\cdot), \cos(\pi\cdot)\}$, which are chosen to build non-linear correlations in the feature space, such as periodicity with sine and cosine operators or quadratic correlations with the square operator. The set of binary operators is $O_2 = \{+, -, \times, \div, |\div|\}$. The binary operators capture complex interactions between the primary features arising from the interplay of different physical phenomena. Physical consistency is maintained by applying addition and subtraction operators only to properties that have the same physical dimension (e.g., feature A with units of eV cannot be added to feature B with units of Å). At each FE step, we remove features that are redundant [e.g., $(x_1 + x_2)$ and $(x_2 + x_1)$], constant, or diverged to infinity. The identification of repeated features [e.g., $(x_1 + x_2)$ and $(x_2 + x_1)$] and the detection of divergence was done by comparing the values of the features numerically instead of comparing their symbolic composition. The exact sequence in which the FE steps are applied is explained in detail below for each FS method. Standard data normalization is enforced by shifting the mean and variance of input features to zero and one, respectively, for the training set alone before FS, which avoids weighting issues caused by the variable magnitude of features with differing units.[12,13] After FS, the training data of selected features is un-normalized to its original magnitude for FE steps.

C. Feature selection algorithms

In a one-shot FE → FS framework, the primary features are engineered in one recursive step and then are passed to the FS algorithm to arrive at the final predictive model in the form of Eq. (1). Complex functional forms of the features can be achieved with sufficient FE steps, and arbitrary accuracy of the predictive model can also be achieved. Thus, one-shot methods can also achieve high multiplicity in terms of the selected features. However, in practice, the size of the feature space after the FE step is limited by available computational resources, and a key challenge associated with one-shot FE is the fact that the number of features grows double exponentially with the number of FE steps. If we apply m binary operators on p features h times recursively, the resulting feature space will contain $p^{2^h} m^{2^h-1}$ candidates. For instance, applying five binary operators on 15 features recursively three times generates $\sim 2 \times 10^{14}$ features. Given this limitation, iBART instead follows an iterative framework that interleaves the FE and FS steps. Here, we will compare the performance of iBART to various one-shot FS methods, including SISSO for both the M–S and M–D datasets, LASSO + l_0 for the M–S dataset,[12] and Dirichlet–Laplace prior for the M–D dataset,[13] with the optimal settings under available computational resources in this work (one Intel Xeon Gold 6230 CPU @ 2.10 GHz node with 40 cores and 960 GB of memory).

SISSO. SISSO is a method, developed by Ouyang *et al.*,[25] that can efficiently search for meaningful features from over billions of candidates. The robust nature of SISSO is demonstrated in numerous studies.[25,27,32] The algorithm is composed of two parts: (1) sure independence screening (SIS)[39] and (2) sparsifying operator (SO). In a one-shot framework, FE is applied on the primary features with $O_1 \cup O_2$ operators in a recursive manner to obtain a massive feature space. Then, each candidate in the feature space is evaluated by SIS to determine its correlation with the property of interest (y). Then, a set number of features with the highest ranked SIS scores are retained. This screening process is computationally efficient, and thus, the input feature space can include as many as 10^{10} features.[25] After the first SIS step, the set of screened features (i.e., S_{1D}) containing 10^1–10^2 features serves as the input feature set for the SO step. SO represents any sparse regression method, i.e., l_1 norm or l_0 norm regression. The purpose of SO is to determine the optimal nD descriptors from S_{nD}. For instance, the best one feature (1D) that predicts the property of interest with the lowest prediction error [e.g., as measured by the root mean square error (RMSE)] can be selected by SO from S_{1D} with simple regression. The SIS–SO process can be repeated multiple times until the prediction error is lower than a user-defined threshold, and the nth set of SIS-selected features is denoted as S_{nD}. Two aspects of the algorithm can be modified as the SIS–SO process is repeated. First, while the first SIS process computes the correlation between the features and y, the nth

SIS evaluates the correlation between the features and the prediction error from the previous run [i.e., $y - \hat{y}_{(n-1)D}$, where $\hat{y}_{(n-1)D}$ is the predicted value from the model of the (n − 1)-th SIS–SO step]. Second, the input space for SO is the union of all previous SIS-selected features (i.e., $S_{1D} \cup S_{2D} \cup \cdots \cup S_{nD}$). In this strategy, the first iteration of SISSO identifies the dominant feature from the S_{1D} set, and the following S_{nD} ($n \geq 2$) sets reduce the gap between y and the current model prediction. We refer to the readers to Ouyang *et al.*[25] for a comprehensive overview of the SISSO algorithm.

Here, multiple parameters are specified when implementing SISSO for the M–S and M–D datasets. The $O_1 \cup O_2$ operator set was applied to the primary features twice to generate ~5.5×10^7 features. The upper bound and lower bound of the absolute value of the descriptors were 1×10^6 and 1×10^{-6}, respectively. The size of the SIS-selected subspace was 40, l_0 norm regression was applied in the SO step, and the maximum dimension for the l_0 norm regression step was D = 5. These parameters settings were chosen because they generate the most candidate features that could be managed with computational resources available to us.

LASSO + l_0 and Dirichlet–Laplace prior. In addition to SISSO, we also compare iBART to the methods we originally employed on the M–S and M–D datasets, which are LASSO + l_0[12] and Dirichlet–Laplace prior,[13] respectively. To ensure a fair comparison with the published results, we apply the same datasets, FE procedures, and FS algorithms as those applied in the original literature since significant effort was put into optimizing settings in LASSO + l_0 and Dirichlet–Laplace prior in our previous works; the only difference here is that we now assess average model performance over multiple trials with different random data separations into training and validation data. The LASSO + l_0 method, which was introduced by Scheffler *et al.*,[23] is implemented in MATLAB,[12] where the shrinkage coefficient of LASSO, λ, is chosen to identify the top 75 features from ~300 000 candidate features, and l_0 norm regression is used to refine the model to the desired final dimension. Dirichlet–Laplace prior is a Bayesian sparse regression algorithm,[40] which uses the Dirichlet–Laplace distribution on each coefficient, updates the distribution with a Gibbs sampler, and then selects meaningful features from the posterior distribution. The entire feature space is pre-screened by the Pearson coefficient of correlation to reduce the amount of candidate features to one thousand. Dirichlet–Laplace prior is then used to conduct feature selection on the candidates. Once the relevant features are identified in this manner, l_0 norm regression again is used to refine the model to a user-specified final dimension. We refer to the reader to our previous work[13] and the statistics literature[40,41] for more details regarding the Dirichlet–Laplace prior method.

iBART. iBART is an algorithm that iteratively applies FE and FS on the feature space.[33] Simple linear regression can select useful features from a pool of candidates if they are linearly correlated with the property of interest (i.e., y). However, in physical systems the property of interest often is not linearly dependent on the features. Therefore, a FS method designed to identify non-linear correlations is advantageous, which is why we selected BART[34,35] as the principal FS method in the iBART routine.

BART is a sum-of-trees model (see the section titled Method in the supplementary material for a basic introduction to decision tree-based models), which can be expressed as

$$y = \sum_{k=1}^{m} g(x; T_k, M_k) + \varepsilon, \tag{2}$$

where $g(\cdot)$ denotes the regression tree function, T_k denotes the structure of a binary tree, $M_k = \{\mu_{k1}, \mu_{k2}, \ldots, \mu_{kb}\}$ represents the parameter set in T_k having b terminal nodes (i.e., leaves), μ_{kl} represents the parameter on each leaf in T_k, m represents the total number of trees, and ε is the error term assuming a normal distribution with zero mean and variance σ^2. The first step when building a BART model is to determine the structure of each tree (T_k). With the initial root node (depth, $d = 0$), the probability of the node becoming a branch is $\alpha(1+d)^{-\beta}$, where α is a number between zero and one and β is a positive number. Values of $\alpha = 0.95$ and $\beta = 2$ typically are applied in the literature to regularize the size of each tree.[34] Thus, as d increases, the node is more likely to become a leaf instead of a branch because $\alpha(1+d)^{-\beta} \to 0$, which terminates the growth of the tree. Once the trees are built, each branch node is assigned a feature and a splitting value, and the leaf nodes are assigned a value, μ_{kl}. A uniform distribution is used for each feature to determine which feature is used in each node, where the splitting values are chosen from the "available" values from another uniform distribution. We note that a splitting value is "available" if the value can split the samples into two categories.[34,42] For μ_{kl}, a normal distribution with zero mean and $\frac{1}{16m}$ variance generally is used as the prior distribution, since the normal distribution is computationally efficient and the variance yields 95% prior probability for y in [−0.5, 0.5] on the normalized scale.[34] Finally, the variance of the error term, σ^2, has an inverse chi-square distribution as the prior distribution, with the hyperparameters for the inverse chi-square distribution chosen from the standard deviation of the prediction error using a linear regression of all features on y. The choice of prior distribution for each parameter is discussed in the original BART literature[34] with comprehensive mathematical derivations found therein.

After the BART model is established from the prior distributions on T_k, M_k, μ_{kl}, and ε, BART implements Bayesian backfitting Markov chain Monte Carlo (MCMC) with the input data (i.e., the input features and y) to obtain the posterior distribution. The number of MCMC steps is critical for obtaining a stable posterior distribution, and this metric dominates the performance of BART models: too few MCMC steps lead to an inaccurate and unstable model, while too many steps waste computational resources. In this work, we use 15 000 MCMC iterations, where the first 10 000 iterations are removed and the last 5000 samples are analyzed for the posterior distribution to determine the final values of each parameter in the tree model. The number of MCMC steps was tested to confirm that the FS results are converged.

BART can be adapted readily as a powerful FS algorithm.[34,35] The BART FS algorithm permutes the prediction target, y, into y' by switching labels and deriving a BART model for each set of y'. Subsequently, the usage frequency of each feature in all BART models is collected. A higher usage frequency implies that the feature is more important statistically for predicting y and, thus, should be selected. Two key factors controlling the effectiveness of FS via BART are the number of MCMC steps and the number of trees, m, which were investigated to determine the proper value while the setting of other parameters follows previous literature.[34,35] If m is large, the total number of nodes is large and irrelevant features can be

selected as false positives. As m decreases, the number of vacancies becomes rare and the model is restricted to only the most meaningful features. Here, we tested $m = 20, 50, 100$ and did not find a significant difference in FS results. Thus, we assign $m = 20$ in the BART models, which is in line with heuristic values recommended in the literature.[34,35] For more information, we refer to the readers to Ref. 34 for BART model construction and Ref. 35 for FS with BART. In addition, Bleich *et al.*[35] demonstrated both that BART can rapidly identify true features in linearly dependent datasets and that BART surpasses random forests and dynamic tree methods when handling non-linearly dependent datasets. Thus, we incorporate BART as the FS method to screen through the non-linearly correlated features that are present in our system.

In the iBART framework (Fig. 2), BART[34] (implemented in R[35]) first is applied to identify the meaningful features from the primary feature set. Subsequently, the O_2 FE operation is employed on the selected features to generate the next batch of candidates. The engineered feature space then is reduced by another BART selection step and then expanded by the O_1 FE operation. The procedure of $BART \rightarrow O_2 \rightarrow BART \rightarrow O_1 \rightarrow BART \rightarrow O_2 \rightarrow BART \rightarrow O_1 \rightarrow \ldots$ can be repeated until a user-defined convergence criterion is achieved. As the FE–FS steps iterate, the dependence between the features and y becomes more linear due to the presence of complex functional forms in the engineered features. Given the increased linear correlation with y after the FE–FS iterations, we implement LASSO[36] (also in R[45]) to conduct the final refining FS step in iBART, where λ is tuned via a tenfold cross validation. For determining λ, we separate the entire dataset into ten subsets, use nine subsets to train the coefficients, and evaluate the prediction error with the remaining one. Each subset is treated as a test set once, and λ with the lowest average prediction error is selected (detailed information can be found in the supplementary material of Ref. 13).

Since we aim to have a fair comparison between iBART and the one-shot approaches, we fix the number of FE steps in iBART to the following sequence: $BART \rightarrow O_2 \rightarrow BART \rightarrow O_1 \rightarrow BART \rightarrow O_2 \rightarrow LASSO \rightarrow l_0$. This yields final features that have a similar level of complexity compared to those derived by the one-shot FE step used for the other FS methods. For SISSO, the $O_1 \cup O_2$ operator set is applied to the primary features twice as mentioned above. For LASSO + l_0 and Dirichlet–Laplace, we use the same FE steps stated in the original literature.[12,13] We note that the order of the FE steps is flexible in iBART; however, we suggest applying O_2 first because it increases feature complexity by introducing interactions between the features. In contrast, O_1 usually derives features that are highly correlated with previous features (i.e., x_1^3 derived from x_1) and does not introduce complexity as quickly.

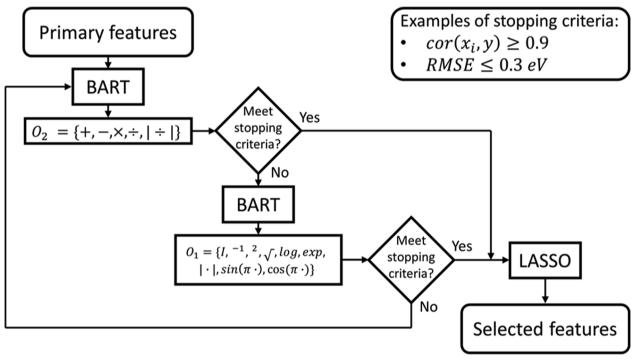

FIG. 2. Flowchart of the iBART algorithm.

D. Evaluation of feature selection methods

For the M-S dataset, we randomly separate the data into a training set (90%) and validation set (10%) for 50 independent trials, while for the M–D dataset, the training-to-validation ratio is 80% to 20% for 50 independent trials. Training sets are used to select the features and to construct the predictive models. Since each FS method will select a different number of features, we use l_0 norm regression to identify the best set of features with a specified model dimension (nD), where n is the number of features used in the final model. We assess the performance of the FS methods using models derived to have the same dimension, which ensures that each model has the same number of free fitting parameters. The root mean squared error (RMSE) of the validation sets is computed to evaluate the final performance. The summary of validation RMSE in 50 random data separations was used as the metric to compare the effectiveness of the FS methods regardless of the training data. The mean of validation RMSE represents the average performance, while the distribution shows the sensitivity of the FS algorithm toward the training data. We note that the results presented herein for LASSO + l_0 on the M–S dataset and Dirichlet–Laplace prior on the M–D dataset are re-derived in this work because here we repeated 50 independent trials with different random separations of training and validation data. The FS methods are compared using the same data in each trial.

III. RESULTS AND DISCUSSION

A. Predicting metal–support binding energies

Datasets from our previous work[12,13] studying metal–support interactions were used to assess the performance of iBART. Metal particles attached on oxide supports are ubiquitous as heterogeneous catalysts in chemical industries. The interaction strength between metals and oxides, measured by the metal binding energy to the support, largely determines the migration barrier of metal adatoms along the surface.[28] Weak binding strength tends to lead to aggregation of metal clusters and deactivation of the catalysts. The key to synthesizing stable cluster morphologies is finding ways to strengthen the metal–support interaction, such as adding dopants that modulate charge transfer [Fig. 3(a)]. In addition to influencing metal particle size distribution, metal–support interactions also alter the reactivity of the metal by changing the metal's oxidation state.[12,29] As such, physical models that can predict metal–support interactions are valuable for designing catalysts with optimal activity and stability.

The M–S dataset from Ref. 12 contains binding energies of 13 single transition metal atoms adsorbed on seven oxide supports, ranging from irreducible MgO to highly reducible TbO_2. The M–D dataset from Ref. 13 contains binding energies of 19 single transition metals on the MgO(100) surface with Al, B, Li, or Na substitutionally doped on a surface Mg lattice site. Here, we apply the same

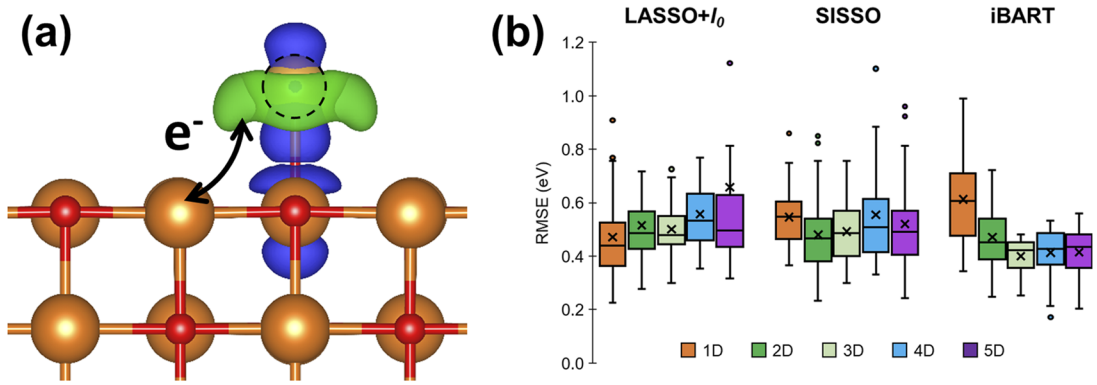

FIG. 3. (a) Isostructural charge density difference plot of Au supported on MgO(100). Blue represents electron density depletion and green represents accumulation. The isosurface level is ±0.0035 *e* Bohr^{-3}. (b) Box plots of validation RMSE for FS-derived models trained on 50 random separations of the M–S dataset under a user-specified model dimension (*n*D). The cross, upper whisker, upper bound, center line, lower bound, and lower whisker represent mean, maximum, 75th percentile, median, 25th percentile, and minimum, respectively. The outliers for box plots are marked as circles, and we note that some outliers for LASSO + l_0 and SISSO are not visible in the plot range. Data in (b) is from Ref. 33.

primary features as in the original studies, which were taken from several available databases. In the M–S set, we have $n = 91$ binding energies and $p = 59$ primary features consisting of atomic properties and oxide properties. This dataset is used herein to evaluate the performance of SISSO and iBART, which we compare against the LASSO + l_0 method used in the original paper. Here, we implement LASSO + l_0 using the original MATLAB code published in Ref. 12. Identical datasets are used to compare LASSO + l_0 with SISSO and iBART. We note that this comparison between LASSO + l_0, SISSO, and iBART using the M–S dataset also is reported in our separate work[33] as a demonstration of a real data application for iBART. A deeper physical interpretation of the results is provided in Sec. III B. We refer to the readers to the original papers[12,13] for further details regarding the full computational details and primary feature data.

We evaluate the SL methods by comparing their average training and validation RMSE over 50 trials, as shown in Fig. 3(b) and in the supplementary material [Fig. 1(a)]. The training error decreases as the dimension of the model increases, as expected. However, at higher model dimensionality, LASSO + l_0 and SISSO both yield models with increasing RMSE. The distribution of RMSE values over the 50 trials also becomes very wide for high model dimensions (i.e., 4D and 5D). These results suggest that there is an early onset of overfitting as the number of features in the model is increased. In contrast, the validation error of iBART reaches a plateau after 3D. At 4D and 5D, the distribution of iBART's validation error becomes slightly wider, which also indicates overfitting in higher dimensional models. Nevertheless, the range in validation error of iBART is narrower at both 4D and 5D compared to LASSO + l_0 and SISSO.

Given the robust performance of iBART on the M–S dataset, we next tested the performance of iBART on the M–D dataset. In these data, a dopant [Fig. 4(a)] is introduced in the oxide surface to tune the metal–support interaction.[46,47] MgO(100) was chosen as the oxide support because it binds to metals weakly, and, thus, the change in metal binding energy induced by the dopants can be observed more apparently. The predicted property from the SL-derived model is now the change in metal binding energy induced by surface dopants, as opposed to the absolute binding energies contained in the M–S dataset. Here, we compare SISSO and iBART with Dirichlet–Laplace prior, which was demonstrated to be more effective than LASSO in the original study.[13] For all FS methods applied on the M–D dataset, the training and validation errors decrease with higher model dimension, as shown in Fig. 4(b) and in the supplementary material [Fig. 1(a)]. SISSO and iBART yield more accurate models than the Dirichlet–Laplace prior method. We found that iBART can derive models with validation error comparable to SISSO with a narrower distribution of error over 50 trials of random training/validation separations. This demonstrates that iBART performs more consistently and is more proficient at creating generalized models.

In addition to predictive power, we aim to demonstrate the interpretability of selected features from iBART and other FS tools in Sec. III B. Here, we describe how we chose representative features for further interpretation. For the M–S dataset, the metal binding energy correlates well with the 1D feature (>0.9 correlation; see Figs. 2 and 3 and Tables I and II of the supplementary material). Therefore, we will further investigate the functional forms and the primary features in the 1D features for all FS methods and the best 2D feature derived from iBART. However, for the M–D dataset, the 1D feature does not correlate with the change in metal binding energy (see Figs. 4–6 and Tables III–V of the supplementary material), which suggests that 2D and 3D features (see Figs. 7–8 of the supplementary material) are required to describe the charge transfer between the doped-surface and the adsorbed metal. We further decompose iBART's 2D and 3D features in Sec. III B 3 to understand which features can describe the bi-directional charge transfer in M–D dataset.

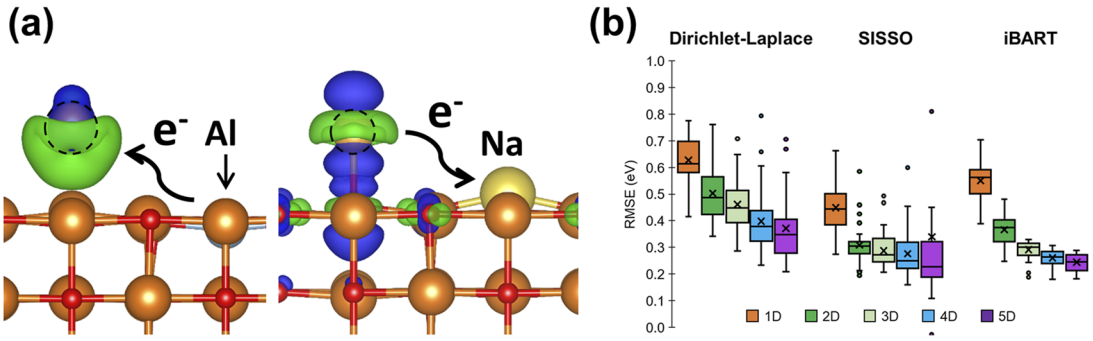

FIG. 4. (a) Isostructural charge density difference plot of Au on Al-doped MgO(100) and Au on Na-doped MgO(100). Blue represents electron density depletion and green represents accumulation. The isosurface level is $\pm 0.0035e$ Bohr^{-3}. (b) Box plots of validation RMSE for Dirichlet–Laplace prior, SISSO, and iBART models trained on 50 random separations of the M–D dataset with user-specified dimension (nD). The cross, upper whisker, upper bound, center line, lower bound, and lower whisker represent mean, maximum, 75th percentile, median, 25th percentile, and minimum, respectively. The outliers for box plots are marked as circles, and we note that some outliers for Dirichlet–Laplace and SISSO are not visible in the plot range.

B. Interpreting physical descriptors derived with iBART

1. Variation in derived features for metal–oxide interactions

The advantage of FS approaches over nonparametric regression is the interpretability of the predictive model. Here, we take interpretability to mean the ability to apply domain knowledge, post-hoc, to understand the physical rules that are embedded in the functional forms of the selected features. For example, in our original study,[12] LASSO + l_0 built a model containing the descriptor $\frac{\Delta H^m_{ox,f}}{O_{vac}}$, where $\Delta H^m_{ox,f}$ is the oxide formation enthalpy of the adsorbed metal and O_{vac} is the oxygen vacancy formation energy of the oxide surface. This functional form is an intuitive measure of charge transfer from the metal to the support. $\Delta H^m_{ox,f}$ has been proposed previously by Campbell and Sellers[48] as a descriptor for predicting metal adhesion on oxides, as it captures the ability of the metal to form bonds with oxygen atoms.[29,37] More generally, as shown by the SL-derived models, $\Delta H^m_{ox,f}$ quantifies the ability of the metal to donate electrons when bonding to an electrophilic element (such as oxygen). O_{vac} describes the ability of the surface to receive electrons, which occurs when an oxygen vacancy is formed. In this work, the 1D features identified from the M–S dataset by each FS method (i.e., LASSO + l_0, SISSO, and iBART) are correlated with the metal's binding energy (i.e., >0.9 Pearson coefficient). Particularly, SISSO derives the same feature each time over 50 random training set separations: $\Delta H^m_{ox,f} \times Z^s \times IE^s_3$, where Z^s is the atomic number of the parent metal in the oxide and IE^s_3 is the third ionization energy of the parent metal in oxide. In contrast, LASSO + l_0 and iBART selected 4 and 22 different features, respectively (supplementary material, Figs. 2 and 3). All 1D features from LASSO + l_0 include $\frac{\Delta H^m_{ox,f}}{O_{vac}}$ (supplementary material, Table I), and the correlation between each feature is >0.98, indicating that they have similar explanatory power for predicting the metal binding energy (supplementary material, Fig. 2). The 1D features from iBART mix $\Delta H^m_{ox,f}$ with various other features (supplementary material Table II), which implies that iBART is sensitive to the training data when constructing functional forms from parent features. However, as we analyzed the selected features from iBART (supplementary material, Fig. 3), we found that they are highly correlated with each other (i.e., the absolute value of correlation between each feature is >0.90), demonstrating that they are describing the same physical phenomenon in different functional forms.

The variation in selected features from iBART is attributed to the nature of BART and the iterative framework applied here. Compared to iBART, one-shot methods, such as LASSO + l_0 and SISSO, generate the entire feature space for FS at the onset. In theory, if computational storage space is unlimited, the functional form of the features can be as complex as desired, and the selected features can form a rich menu of descriptors in various mathematical functions with different primary features. However, the availability of computational resources is a bottleneck in practice, thus limiting the complexity attainable in the one-shot methods. Therefore, one-shot methods are more likely to select the same features, as they are less sensitive to the specific sampling in the training data. In contrast, iBART is sensitive to the training data, as we observe that differences in the training data sampling significantly impact the path taken during FE. As a result, iBART derives a richer menu of features than one-shot approaches. We note that the difference in variation of selected features is observed from our two datasets, and in principle, if computational resources are unlimited, one-shot approaches can also derive a diverse set of descriptors. In addition to deriving diverse features, another key advantage of iBART is its computational efficiency, which will be discussed in Sec. III C.

There is an inevitable trade-off between model accuracy and interpretability. Descriptors with simple functional forms will capture the general physics and can be interpreted easily, but they usually suffer reduced accuracy. In contrast, descriptors with complex functional forms usually yield more accurate predictions, but the user may not be able to understand the underlying physics in the selected descriptors. iBART still suffers from this trade-off, but an advantage of iBART is its ability to produce a rich menu of derived features that take unique functional forms but describe the same

physics. This attribute raises the possibility that iBART can generate apparent contradictions, as different descriptors can be selected to describe the same physics. However, for a given training set, iBART will always build features that linearly correlate with the property of interest. Thus, if two unique descriptors are both correlated with the property of interest, they must also be correlated with each other. This can allow the user to identify physical relationships between the descriptors that were not known previously.[49] iBART builds complexity as the algorithm progress, and thus, it is more flexible and can explore more possible functional forms.

An example of this attribute of iBART is provided here, where we develop a connection between oxygen vacancy formation energy (O_{vac}) and ionization energies of parent metals in the oxide surface. As shown in Table II of the supplementary material, we can derive multiple replacements for the O_{vac} feature that provide the same information. For example, we find high correlations if we compare $\left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right|$ (feature x.7 in Table II of the supplementary material) with $\Delta H^m_{ox,f} \times EA^s \times IE^s_4$, $\Delta H^m_{ox,f} \times (IE^s_4 + EA^s) \times IE^s_4$, and $\Delta H^m_{ox,f} \times Z^s \times (IE^s_3)^2$ (features x.8, x.9, and x.10 in Table II of the supplementary material, respectively), where EA^s is the electron affinity of the parent metal in the oxide and IE^s_4 is the fourth ionization energy of the parent metal in the oxide. Although O_{vac} is an effective descriptor for the metal binding energy, this property is obtained from a density functional theory (DFT) calculation. If the goal is to reduce the number of required DFT calculations, then it would be advantageous to use equivalent properties available in experimental databases. Thus, the flexibility of iBART allows us to select primary features that are preferable based on availability or accuracy (i.e., selecting more exact experimental descriptors over computed descriptors).

2. Decomposing metal-oxide interactions between single metal atoms and oxide supports

The model for the M–S dataset can be further improved by increasing the dimension, i.e., including more features in the predictive model. The best 2D model derived from iBART includes $\left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right| \times \frac{1}{\phi^s}$ and $\frac{H^m_f \times EA^s}{NVal^m}$, where ϕ^s is the Miedema parameter representing the chemical potential of the electrons in the parent metal in the oxide, H^m_f is the oxide formation enthalpy in its most stable state of the adsorbed metal, and $NVal^m$ is the number of valence electrons of the adsorbed metal. The first feature ($\left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right| \times \frac{1}{\phi^s}$) correlates with $\left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right|$ (i.e., $cor\left(\left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right|, \left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right| \times \frac{1}{\phi^s}\right) = 0.99$), which captures the metal–oxygen interactions and correlates well with the metal binding energy [Fig. 5(a)] for all oxides, as expected from our previous discussion. Data for most oxides, except MgO(100) and ZnO(100), cluster on a trend line (see slopes for each oxide in Table VI of the supplementary material). This implies that $\left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right| \times \frac{1}{\phi^s}$ is a suitable descriptor, in general, but does not describe the irreducible oxides and reducible oxides in the same way. The second descriptor ($\frac{H^m_f \times EA^s}{NVal^m}$) has a negative correlation for MgO(100) and ZnO(100) but a positive correlation for the other surfaces [Fig. 5(b)]. The sign of this correlation is controlled by the EA^s term in $\frac{H^m_f \times EA^s}{NVal^m}$. For Mg and Zn, EA^s is negative, while for other parent metals, EA^s is positive. From a chemical point of view, this implies that for irreducible oxides, the parent metals do not favorably accept additional electrons from the adsorbed metal and, therefore, will be inert in the charge transfer process. In contrast, for reducible oxides, electrons from adsorbed metals will interact with not only the surface oxygen but adjacent parent metal atoms, which enhances charge transfer and strengthens the binding energy.

Analysis of the 2D model derived from iBART provides important information for metal–oxide interactions. Metal binding strength is determined by the charge transfer between the adsorbed metal, surface oxygens, and parent metals in the oxide. Here, we deconstruct the metal–oxide interactions into (1) redox interactions between the adsorbed metal and surface oxygens (i.e., explicit electron transfers), (2) bonding interactions between the adsorbed metal and surface oxygens (i.e., atomic orbital rehybridizations), (3) redox interactions between the adsorbed metal and the surface parent metals, and (4) bonding interactions between the adsorbed metal and the surface parent metals. $\left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right|$ is generally the descriptor that correlates most with the binding energy because it captures the metal–oxygen interactions for both redox and bonding interactions. To further improve the model, descriptors for the charge transfer between the adsorbed metal and parent metal, i.e., metal–metal redox interactions and metal–metal bonding interactions, are critical. Such interactions are captured by the term $\frac{H^m_f \times EA^s}{NVal^m}$ using H^m_f and EA^s to describe the charge transfer ability of the adsorbed metal and the surface parent metal, respectively. We conducted Bader analysis, crystal orbital Hamilton populations (COHP) analysis, and generated charge density different plots using DFT to illustrate the charge transfer between the adsorbed metal and the oxide surfaces (see the section titled Method in the supplementary material for computational settings). In particular, Bader analysis reveals the charge in each atom and their redox interactions, while COHP analysis determines whether direct bonds are formed between two atoms. Ag and V were chosen to represent low- and high-oxyphilic metals, respectively, while MgO(100) and CeO_2(111) were chosen to represent irreducible and reducible oxides, respectively (supplementary material, Fig. 9).

On MgO(100), Ag and V form an anti-bonding interaction with attached oxygens [supplementary material, Figs. 10(a) and 10(c)], and no obvious redox interactions are present (supplementary material, Table VII). In addition, both redox and bonding interactions between Ag or V and adjacent Mg atoms are negligible, indicating weak metal–metal interactions due to the negative EA of Mg. Figure 5(d) demonstrates the charge transfer between the adsorbed V and MgO(100) surface, where no V–Mg bond is observed in the COHP analysis. On CeO_2(111), stronger Ag–O interactions are evident in the charge density difference plot [Fig. 5(e)]. COHP analysis indicates the formation of Ag–O bonds with surface oxygens [supplementary material, Figs. 11(a)–11(d)]. Additionally, redox interactions between Ag and Ce atoms are visible in the Bader analysis, where Ag atoms are clearly oxidized and Ce atoms are reduced (supplementary material, Table VII). This is due to the higher EA of Ce, illustrating how the $\frac{H^m_f \times EA^s}{NVal^m}$ feature derived by iBART captures redox interactions. However, no Ag–Ce bonds are evident in the COHP analyses [supplementary material, Figs. 11(e)–11(h)].

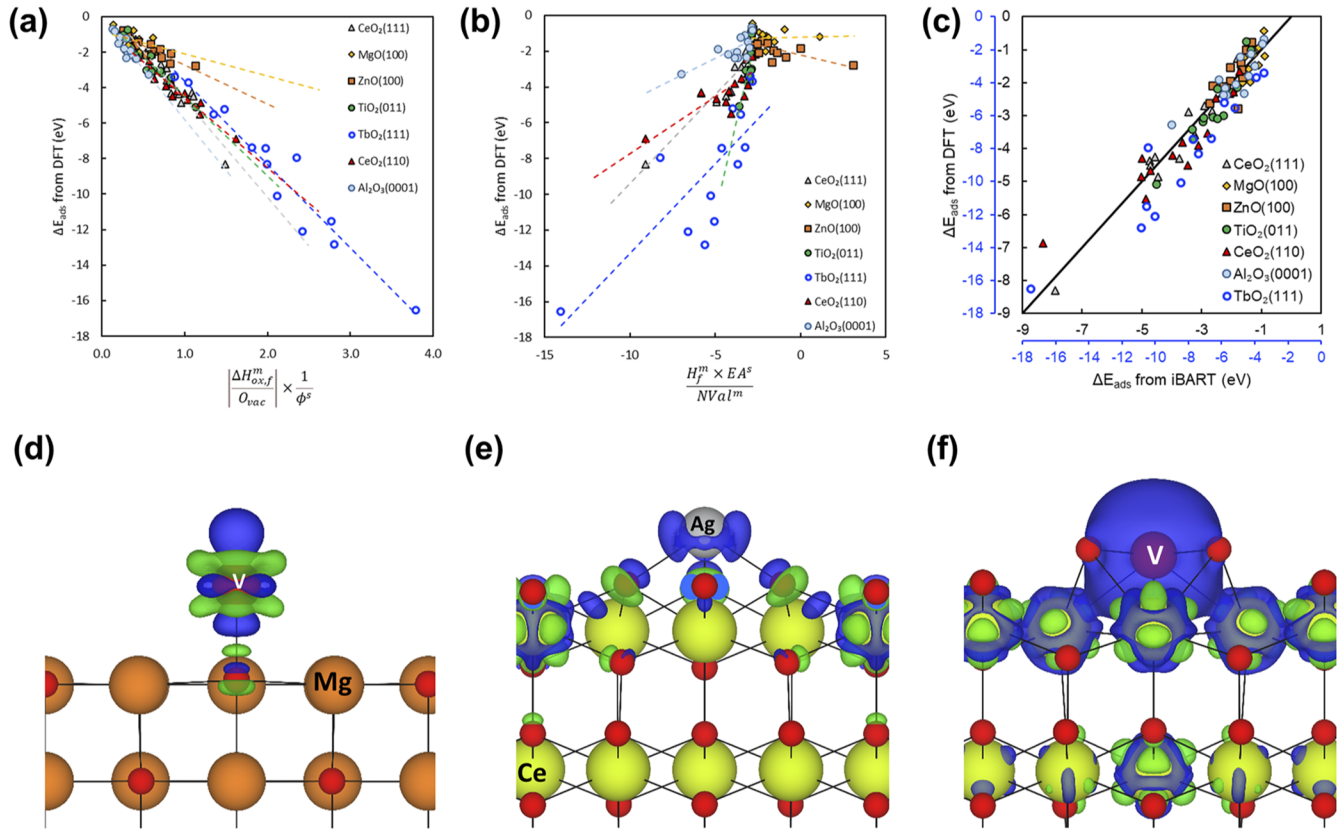

FIG. 5. Correlation between metal binding energy from the M–S dataset with (a) and (b) individual features and (c) a 2D model ($\Delta E_{ads} = -3.29 \times \left|\frac{\Delta H^m_{ox,f}}{O_{vac}}\right| \times \frac{1}{\phi^s} + 2.43 \times \frac{H^m_f \times EA^s}{NVal^m} - 0.47$) derived from iBART. The lower panels show charge density difference plots for (d) V/MgO(100), (e) Ag/CeO_2(111), and (f) V/CeO_2(111). Blue represents electron density depletion and green represents accumulation. The isosurface level is ±0.005*e* Bohr^{-3}. The red, dark red, orange, green, and gray spheres represent O, V, Mg, Ce, and Ag, respectively.

Thus, strong Ag–O interactions and the redox interactions between Ag and Ce result in stronger binding on the CeO_2(111) surface (−1.73 eV) compared to the MgO(100) surface (−0.44 eV). For V adsorbed on CeO_2(111), similar to Ag/CeO_2(111), V–O bonds were observed in the COHP analysis [supplementary material, Figs. 12(a)–12(d)] and charge density difference plots [Fig. 5(f)]. Strong redox interactions between V and Ce atoms are evident in the Bader analysis (supplementary material, Table VII). Compared with Ag, V is more oxyphilic, resulting in stronger V–O bonds and enhanced redox interactions between V and Ce. Moreover, direct V–Ce bonds are demonstrated in the COHP analysis [supplementary material, Fig. 12(e)–12(h)] and charge different density plots [Fig. 5(f)]. Therefore, V exhibits a strong interaction with surface oxygens, as well as strong redox and bonding interactions with surface Ce atoms, leading to much stronger binding energy with the CeO_2(111) surface (−8.30 eV).

These electronic analyses show that several distinct interactions between adsorbed transition metals and oxide surfaces dominate the overall metal binding energy. In terms of iBART derived models, the 1D model captures the general trend [Fig. 5(a)] with a descriptor predicting the strength of metal–oxygen interactions. An additional descriptor added in the 2D model ($\frac{H^m_f \times EA^s}{NVal^m}$) measures the strength of direct metal–metal interactions, i.e., through both redox and bonding interactions, which differentiate irreducible oxides from the reducible oxides. We note that our previous work also discovered descriptors responsible for metal–metal interactions with LASSO + l_0, but this feature showed up only in 3D or higher dimensional models.[12]

3. Bi-directional charge transfer between metal atoms and doped-oxides

The flexibility of models derived by iBART also is advantageous when deriving models on datasets that contain several categorical distinctions, which is the case for the M–D dataset. In M–D dataset, dopants were introduced to the surface that promoted electron transfer either to the metal from the surface or vice versa. The descriptors derived for the M–S dataset, e.g., $\Delta H^m_{ox,f}$, fail to describe the M–D data because they only describe electron transfer from the metal to the surface. A more complex model is

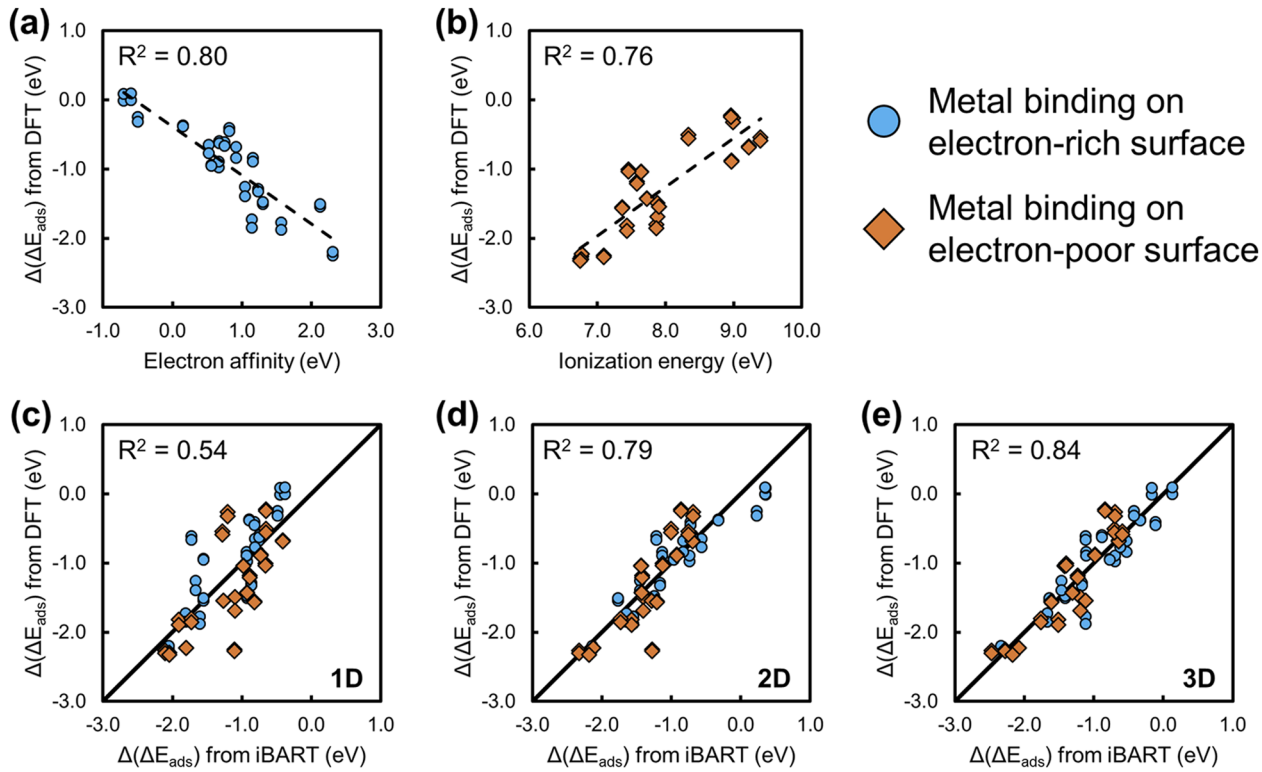

FIG. 6. Correlation between change in metal binding energy difference from the M–D dataset with (a) electron affinity and (b) ionization energy of the adsorbed metal and (c)–(e) 1D–3D features derived from iBART, listed in Table I. Data in (a) and (b) are from Ref. 13.

TABLE I. Predicting equations composed of features constructed with iBART and their prediction error on the full M–D dataset.

	Predicting equation	RMSE (eV)
1D	$1.66 \times \log\left(\left\lvert\log\left(NVal^m \times NVal^d\right) \times \log\left(\frac{EN_P^m}{NVal^d}\right)\right\rvert\right) - 1.64$	0.45
2D	$-7.88 \times \left\lvert IE_1^m \times NVal^d \times \log\left(\frac{NVal^d}{EN_P^m}\right)\right\rvert^{-1}$ $-0.69 \times \exp\left(\frac{EA^m}{\exp\left(\frac{EA^m}{NVal^d}\right)}\right) + 1.15$	0.31
3D	$0.17 \times H_f^m \times \log\left(\frac{EN_P^m}{NVal^d}\right)$ $-29.85 \times \left(IE_1^m\right)^{-1}$ $-0.22 \times \frac{EA^m}{\log\left(\frac{EN_P^m}{NVal^d}\right)} + 3.16$	0.27

$NVal^m$: number of valence electrons of adsorbed metal

$NVal^d$: number of valence electrons of surface dopant

EN_P^m: Pauling's electronegativity of adsorbed metal

IE_1^m: first ionization energy of adsorbed metal

EA^m: electron affinity of adsorbed metal

H_f^m: oxide formation enthalpy of the adsorbed metal's most stable oxide

needed to capture the bi-directional nature of charge transfer in the M–D dataset. In previous work,[13] we found that both EA^m and IE_1^m were needed to separately describe the electron-rich and electron-poor surfaces, respectively, as demonstrated in Figs. 6(a) and 6(b) (IE_1^m is the first ionization energy of the adsorbed metal). In Table I, we found that $\log\left(\frac{EN_p^m}{NVal^d}\right)$ is repeatedly selected throughout the 1D to 3D predictive models from iBART. As illustrated in Fig. 13(a) of the supplementary material, this feature distinguishes electron-rich surfaces from electron-poor surfaces by using $NVal^d$ as a categorical tag. Since $NVal^d = 3$ for electron-rich surfaces and $NVal^d = 1$ for electron-poor surfaces, the magnitude of $\log\left(\frac{EN_p^m}{NVal^d}\right)$ dominates for electron-poor surfaces and diminishes for electron-rich surfaces, leading to two clusters shown in Fig. 13(a) of the supplementary material. When $\log\left(\frac{EN_p^m}{NVal^d}\right)$ is multiplied by $\log\left(NVal^m \times NVal^d\right)$, we see that the clusters merge and a linear correlation emerges [supplementary material, Fig. 13(b)]. Including an additional $\log(\cdot)$ operator [supplementary material, Fig. 13(c)] and fitting coefficients leads to the final predictive model in Fig. 6(c). However, scatter can be observed in the parity plot of the 1D model, which implies that a higher dimensional model is necessary for a complete description of bi-directional charge transfer. For 2D features, $\left|IE_1^m \times NVal^d \times \log\left(\frac{NVal^d}{EN_p^m}\right)\right|^{-1}$ and $\exp\left(\frac{EA^m}{\exp\left(\frac{EA^m}{NVal^d}\right)}\right)$ were selected, and we found a significant decrease in prediction error [Fig. 6(d)]. Interestingly, one feature includes IE_1^m, while the other is composed with EA^m, matching our prior knowledge that IE_1^m and EA^m are essential for electron-poor and electron-rich surfaces, respectively. For the electron-rich surface, the charge will be transferred from the support to the adsorbed metal, and thus, a

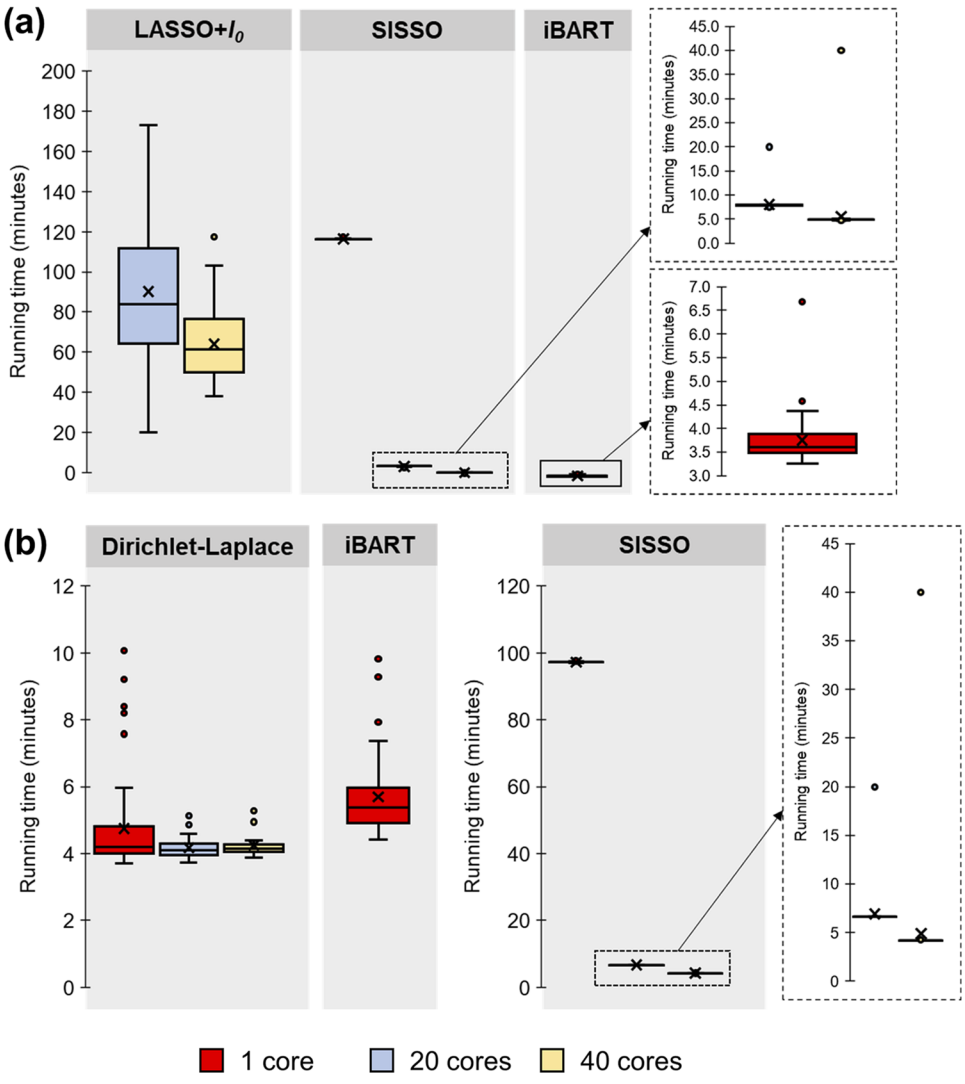

FIG. 7. Box plots of runtime for each method across 50 random partitions in using 1, 20, and 40 CPU cores for the (a) M–S and (b) M–D dataset. The cross, upper whisker, upper bound, center line, lower bound, and lower whisker represent mean, maximum, 75th percentile, median, 25th percentile, and minimum, respectively. Data in (a) are from Ref. 33.

descriptor predicting ability of the adsorbed metal to accept charge, i.e., $\exp\left(\frac{EA^m}{\exp\left(\frac{EA^m}{NVal^d}\right)}\right)$, dominates the predictive model (supplementary material, Table VIII). For electron-poor surfaces, the other feature, $\left|IE_1^m \times NVal^d \times \log\left(\frac{NVal^d}{EN_P^m}\right)\right|^{-1}$, dominates the prediction by describing the charge donation ability of the adsorbed metal. For the 3D features [Fig. 6(e)], we found similar functional forms that use variants of $\log\left(\frac{EN_P^m}{NVal^d}\right)$ as a categorical tag to make classifications in the final model.

The bi-directional charge transfer in the M–D dataset was investigated using electronic analysis in our previous work.[13] Here, no assumption and physical information, except training data, were provided in FE–FS procedures, and iBART was able to obtain the data-driven 2D model for describing the bi-directional charge transfer. EA^m and IE_1^m were used to describe the extent and direction of charge transfer, and $NVal^d$ was applied to distinguish electron-rich from electron-poor surfaces. In this work, we use M–S and M–D datasets as the testbed for iBART to not only measure its prediction accuracy but verify if it can recover physical insights from the given training data. The results indicate that physical models can be established by iBART.

C. Computational cost advantage of iBART

In Secs. III A and III B, we demonstrated the capability of iBART to derive meaningful and predictive features. Here, we demonstrate another advantage of iBART, which is its computational efficiency. In Fig. 7, we compare the computational time for LASSO + l_0, SISSO, and iBART to build a model using the M–S dataset as well as the time for Dirichlet–Laplace, SISSO, and iBART to build a model using the M–D dataset. We compare the required wall-clock times using 1, 20, and 40 cores on an Intel Xeon Gold 6230 CPU @ 2.10 GHz node with 40 cores and 960 GB of memory. When 40 cores were not fully utilized, the remaining cores were excluded from other computational tasks. The LASSO + l_0 run did not finish using one core because of its high computational demand in the l_0 norm regression step, and thus, no data for this run are reported in Fig. 7(a). Among the three FS methods in Fig. 7(a), iBART yields the shortest runtime because it reduces the size of the candidate feature space to a maximum of $\sim 10^2$ entries, compared to $>10^5$ entries for LASSO + l_0 and SISSO. SISSO can be efficiently parallelized, but it is still slower using 40 cores than iBART using one core. A similar trend can be found when applying the FS methods on the M–D dataset [Fig. 7(b)]. Dirichlet–Laplace prior has a comparable speed to iBART, since we pre-screen the feature space with the Pearson coefficient of correlation, and as a result, it only needs to conduct FS on 1000 features. However, the latter introduces greater complexity into the feature space and, thus, can derive more meaningful features [Fig. 4(b)]. We note that the parallelization of iBART has not yet been developed in the current version; however, the runtime can be further reduced if parallel computation is implemented. These results show that iBART is computationally efficient, as it can be implemented over the course of minutes using resources available on a typical personal computer.

IV. CONCLUSIONS

In this paper, we demonstrated that a new FS method, iBART, can derive robust physical descriptors for predicting the behavior of complex catalytic materials with reduced computational cost. On the M–S data, iBART recovers the meaningful descriptors found in our previous work using LASSO + l_0 set while only requiring a fraction of the computational resources. On the M–D dataset, we found that the flexibility of iBART allows it to construct complex features that can describe categorical splits in the training data (i.e., electron-rich vs electron-poor surfaces). iBART constructs models that achieve similar accuracy to those derived with SISSO, yet iBART exhibits less variation in validation error over repeated trials with random separations of training and testing data. This demonstrates that iBART can provide consistent and generalized models. Finally, iBART significantly reduces the size of the candidate feature space and, achieves high computational efficiency, thus allowing it to be implemented on a personal computer using only minutes of runtime.

SUPPLEMENTAL MATERIAL

The supplementary material includes additional descriptions of the computational settings for all DFT calculations, supplemental figures, and supplemental tables that are referenced in the main text.

ACKNOWLEDGMENTS

C.-Y.L. and T.P.S. acknowledge the American Chemical Society Petroleum Research Fund (Grant No. ACS PRF# 59759-DNI). S.Y. and M.L. acknowledge the National Institutes of Health (Grant No. T32CA096520). Computational resources were supported, in part, by the Big-Data Private-Cloud Research Cyberinfrastructure MRI-award funded by the NSF under Grant No. CNS-1338099 and by Rice University. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high performance computing (HPC) resources to achieve testing datasets.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

AUTHOR CONTRIBUTIONS

C.-Y.L. and S.Y. contributed equally. M.L. and T.P.S. designed the research; S.Y. developed the iBART framework; C.-Y.L. helped with refining the code and provided domain knowledge for interpretation; and C.-Y.L., S.Y., M.L., and T.P.S. wrote the paper.

DATA AVAILABILITY

M–S and M–D datasets are available in the following online repository: https://github.com/tsenftle/Metal-Oxide-LASSO-lo, Ref. 50, and https://github.com/tsenftle/MgO_SL, Ref. 51, respectively. iBART source code is available in https://github.com/mattsheng/iBART, Ref. 52.

REFERENCES

[1] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, Nature **559**, 547 (2018).
[2] B. R. Goldsmith, J. Esterhuizen, J. X. Liu, C. J. Bartel, and C. Sutton, AIChE J. **64**, 2311 (2018).
[3] K. M. Jablonka, D. Ongari, S. M. Moosavi, and B. Smit, Chem. Rev. **120**(16), 8066–8129 (2020).
[4] C. Chen, Y. Zuo, W. Ye, X. Li, Z. Deng, and S. P. Ong, Adv. Energy Mater. **10**, 1903242 (2020).
[5] T. Toyao, Z. Maeno, S. Takakusagi, T. Kamachi, I. Takigawa, and K.-i. Shimizu, ACS Catal. **10**, 2260 (2020).
[6] Z. W. Ulissi, A. J. Medford, T. Bligaard, and J. K. Nørskov, Nat. Commun. **8**, 14621 (2017).
[7] C. W. Coley, W. H. Green, and K. F. Jensen, Acc. Chem. Res. **51**, 1281 (2018).
[8] B. Sanchez-Lengeling and A. Aspuru-Guzik, Science **361**, 360 (2018).
[9] J. Noh, G. H. Gu, S. Kim, and Y. Jung, Chem. Sci. **11**, 4871 (2020).
[10] K. Tran and Z. W. Ulissi, Nat. Catal. **1**, 696 (2018).
[11] B. Burger, P. M. Maffettone, V. V. Gusev, C. M. Aitchison, Y. Bai, X. Wang, X. Li, B. M. Alston, B. Li, R. Clowes, N. Rankin, B. Harris, R. S. Sprick, and A. I. Cooper, Nature **583**, 237 (2020).
[12] N. J. O'Connor, A. S. M. Jonayat, M. J. Janik, and T. P. Senftle, Nat. Catal. **1**, 531 (2018).
[13] C.-Y. Liu, S. Zhang, D. Martinez, M. Li, and T. P. Senftle, npj Comput. Mater. **6**, 102 (2020).
[14] B. Medasani, A. Gamst, H. Ding, W. Chen, K. A. Persson, M. Asta, A. Canning, and M. Haranczyk, npj Comput. Mater. **2**, 1 (2016).
[15] J. Schmidt, J. Shi, P. Borlido, L. Chen, S. Botti, and M. A. L. Marques, Chem. Mater. **29**, 5090 (2017).
[16] J. A. Esterhuizen, B. R. Goldsmith, and S. Linic, Chem **6**, 3100 (2020).
[17] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, Phys. Rev. Lett. **115**, 205901 (2015).
[18] G. Pilania, J. E. Gubernatis, and T. Lookman, Comput. Mater. Sci. **129**, 156 (2017).
[19] Z. Li, L. E. K. Achenie, and H. Xin, ACS Catal. **10**, 4377 (2020).
[20] J. N. Wei, D. Duvenaud, and A. Aspuru-Guzik, ACS Cent. Sci. **2**, 725 (2016).
[21] A. Nandy, J. Zhu, J. P. Janet, C. Duan, R. B. Getman, and H. J. Kulik, ACS Catal. **9**, 8243 (2019).
[22] K. T. Schütt, M. Gastegger, A. Tkatchenko, K.-R. Müller, and R. J. Maurer, Nat. Commun. **10**, 5024 (2019).
[23] L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl, and M. Scheffler, Phys. Rev. Lett. **114**, 105503 (2015).
[24] P. Pankajakshan, S. Sanyal, O. E. de Noord, I. Bhattacharya, A. Bhattacharyya, and U. Waghmare, Chem. Mater. **29**, 4190 (2017).
[25] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, and L. M. Ghiringhelli, Phys. Rev. Mater. **2**, 083802 (2018).
[26] C. J. Bartel, C. Sutton, B. R. Goldsmith, R. Ouyang, C. B. Musgrave, L. M. Ghiringhelli, and M. Scheffler, Sci. Adv. **5**, eaav0693 (2019).
[27] M. Andersen, S. V. Levchenko, M. Scheffler, and K. Reuter, ACS Catal. **9**, 2752 (2019).
[28] Y.-Q. Su, L. Zhang, Y. Wang, J.-X. Liu, V. Muravev, K. Alexopoulos, I. A. W. Filot, D. G. Vlachos, and E. J. M. Hensen, npj Comput. Mater. **6**, 144 (2020).
[29] M. E. Strayer, T. P. Senftle, J. P. Winterstein, N. M. Vargas-Barbosa, R. Sharma, R. M. Rioux, M. J. Janik, and T. E. Mallouk, J. Am. Chem. Soc. **137**, 16216 (2015).
[30] S. Curtarolo, D. Morgan, K. Persson, J. Rodgers, and G. Ceder, Phys. Rev. Lett. **91**, 135503 (2003).
[31] J. P. Janet and H. J. Kulik, J. Phys. Chem. A **121**, 8939 (2017).
[32] W. Xu, M. Andersen, and K. Reuter, ACS Catal. **11**, 734 (2021).
[33] S. Ye, T. P. Senftle, and M. Li, arXiv:2110.10195 [Stat] (2021).
[34] H. A. Chipman, E. I. George, and R. E. McCulloch, Ann. Appl. Stat. **4**, 266 (2010).
[35] J. Bleich, A. Kapelner, E. I. George, and S. T. Jensen, Ann. Appl. Stat. **8**, 1750 (2014).
[36] R. Tibshirani, J. R. Stat. Soc. Ser. B **58**, 267 (1996).
[37] S. L. Hemmingson and C. T. Campbell, ACS Nano **11**, 1196 (2017).
[38] C. T. Campbell and Z. Mao, ACS Catal. **7**, 8460 (2017).
[39] J. Fan and J. Lv, J. R. Stat. Soc. Ser. B **70**, 849 (2008).
[40] A. Bhattacharya, D. Pati, N. S. Pillai, and D. B. Dunson, J. Am. Stat. Assoc. **110**, 1479 (2015).
[41] H. D. Bondell and B. J. Reich, J. Am. Stat. Assoc. **107**, 1610 (2012).
[42] H. A. Chipman, E. I. George, and R. E. McCulloch, J. Am. Stat. Assoc. **93**, 935 (1998).
[43] A. R. Linero, J. Am. Stat. Assoc. **113**, 626 (2018).
[44] A. R. Linero and Y. Yang, J. R. Stat. Soc. Ser. B **80**, 1087 (2018).
[45] J. H. Friedman, T. Hastie, and R. Tibshirani, J. Stat. Software **33**, 1 (2010).
[46] Y. Cui, C. Stiehler, N. Nilius, and H.-J. Freund, Phys. Rev. B **92**, 075444 (2015).
[47] G. Pacchioni and H.-J. Freund, Chem. Soc. Rev. **47**, 8474 (2018).
[48] C. T. Campbell and J. R. V. Sellers, Faraday Discuss. **162**, 9 (2013).
[49] L. M. Ghiringhelli *et al.* "The ability to reveal such relationships among selected descriptors is a strength of all symbolic-regression methods," New J. Phys. **19**, 023017 (2017).
[50] A. S. M. Jonayat, "Distributed as part of the publication: Interaction trends between single metal atoms and oxide supports identified with density functional theory and statistical learning." (2018). Github. https://github.com/tsenftle/Metal-Oxide-LASSO-lo
[51] C.-Y. Liu and S. Zhang, "Distributed as part of the publication: Using Statistical Learning to Predict Interactions Between Single Metal Atoms and Modified MgO(100) Supports," (2020). Github. https://github.com/tsenftle/MgO_SL
[52] S. Ye and M. Li, "An R package for iterative BART for Variable and Operator Selection with Operator Induced Structure (OIS)," (2022). Github. https://github.com/mattsheng/iBART