AIP Advances

RESEARCH ARTICLE | JANUARY 26 2024

A data-driven approach to predict the saturation magnetization for magnetic 14:2:1 phases from chemical composition

Amit Kumar Choudhary ; Dominic Hohs ; Andreas Jansche ; Timo Bernthaler; Dagmar Goll ; Gerhard Schneider

Check for updates

AIP Advances 14, 015060 (2024)
https://doi.org/10.1063/5.0171922

View Online
Export Citation
CrossMark

08 February 2024 09:36:03

AIP Publishing

AIP Advances
Special Topic: Novel Applications of Focused Ion Beams — Beyond Milling
Submit Today
AIP Publishing

A data-driven approach to predict the saturation magnetization for magnetic 14:2:1 phases from chemical composition

Cite as: AIP Advances 14, 015060 (2024); doi: 10.1063/5.0171922
Submitted: 11 August 2023 • Accepted: 2 January 2024 •
Published Online: 26 January 2024

View Online Export Citation CrossMark

Amit Kumar Choudhary,[1,2,a)] Dominic Hohs,[1,2] Andreas Jansche,[1] Timo Bernthaler,[1] Dagmar Goll,[1] and Gerhard Schneider[1,2]

AFFILIATIONS

[1] Materials Research Institute, Aalen University, Aalen, Germany
[2] Karlsruhe Institute of Technology, Karlsruhe, Germany

[a)] Author to whom correspondence should be addressed: amit.kumar.choudhary@hs-aalen.de

ABSTRACT

14:2:1 phases enable permanent magnets with excellent magnetic properties. From an application viewpoint, saturation polarization, Curie temperature, and anisotropy constant are important parameters for the magnetic 14:2:1 phases. Novel chemical compositions that represent new 14:2:1 phases require especially maximum saturation magnetization values at application-specific operating temperatures to provide maximum values for the remanence and the maximum energy density in permanent magnets. Therefore, accurate knowledge of the saturation magnetization M_s is important. M_s gets affected by chemical composition in a twofold way, with chemical composition significantly influencing both magnetic moments and crystal structure parameters. Therefore, for magnetic 14:2:1 phases, we have developed a regression model with the aim to predict the saturation magnetization in [μ_B/f.u.] at room temperature directly from the chemical composition as input features. The dataset for the training and testing of the model is very diverse, with literature data of 143 unique phases and 55 entries of repeated phases belonging to the ternary, quaternary, quinary, and senary alloy systems. Substitutionally dissolved elements are heavy and light rare earth elements, transition metals, and additional elements. The trained model is a voting regressor model with different weights assigned to four base regressors and has generalized well, resulting in a low mean absolute error of 0.8 [μ_B/f.u.] on the unseen test set of 52 phases. This paper could serve as the basis for developing novel magnetic 14:2:1 phases from chemical composition.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0171922

08 February 2024 09:36:03

I. INTRODUCTION

Sintered magnets based on the intermetallic phase $Fe_{14}Nd_2B$ are currently the strongest class of permanent magnets. Due to their high remanence, coercivity, and maximum energy density, they are widely used in the motors of electrical vehicles or in the generators of wind turbines.

In such sintered magnets, besides the microstructure, the chemical composition of the hard magnetic 14:2:1 phase ($TM_{14}RE_2B$; TM = transition metal, RE = rare earth, B = boron) is primarily responsible for the outstanding magnetic properties. If solubility permits, elemental substitutes may occupy regular TM and/or RE sites in the 14:2:1 crystal lattice.[1–6] A comprehensive understanding of solubility limits and of substitutional and interstitial sites can offer new directions for 14:2:1 compositional design to meet application-specific magnetic properties.

Exceptional intrinsic magnetic properties of the 14:2:1 phase (saturation polarization $J_s > 1.2$ T, magnetocrystalline anisotropy $K_1 > 10^6$ J/m^3) are prerequisites for large remanences and coercivities, respectively. In RE-TM intermetallics, the high saturation polarization J_s and Curie temperature T_c of the transition metals (TM = Fe, Co) are combined with the high magnetocrystalline anisotropy K_1 of the rare-earth metals (RE = Nd, Pr, Sm).

Besides the Curie temperature,[7–9] the saturation magnetization M_s or saturation polarization J_s ($J_s = \mu_0 M_s$, μ_0: vacuum permeability) of the 14:2:1 phase is therefore an important property for applications. The magnetization is usually reported in the literature using the unit [μ_B/f.u.] (μ_B: Bohr magneton, f.u.: formula unit). The unit

© Author(s) 2024

[μ_B/f.u.] can be easily converted into [emu/g] (cgs). For the engineering of devices, the unit Tesla [T] (SI) is common. For the conversion from [emu/g] to [T], mass density is needed.

Due to Hund's rule, there is an antiparallel coupling between the 3d spins of TM and the 4f spins of RE. For compounds with light RE (Pr, Nd, Pm, Sm, Eu, and Gd), the negative coupling of spins leads to a parallel alignment of RE and TM moments and ferromagnetic temperature behavior (larger J_s values). For compounds with heavy RE (Tb, Dy, Ho, Er, Tm, and Yb), the negative coupling of spins leads to an antiparallel alignment of RE and TM moments and ferrimagnetic temperature behavior (smaller J_s values). For RE without the 4f moment (La, Ce, Lu, Y, Th), Fe sublattices are ferromagnetically coupled.[1,5] Furthermore, different magnetic contributions may occur in saturation magnetization: The mean Fe moments in $Fe_{14}RE_2B$ are around 2.2 μ_B/atom, and the mean Co moments in $Co_{14}RE_2B$ are 1.4 μ_B/atom.[1,5]

With the advancement of data-driven approaches, the shift in the paradigm for material design has been widely driven by materials informatics. Over the last decades, there has been a significant rise in the use of machine learning (ML) techniques to predict the material's magnetic properties and accelerate the process of new composition search.[9–13] Concerning saturation magnetization M_s, there are ML-based regression models developed for the prediction of M_s for hard magnetic phases, such as for rare earth lean intermetallic $REA_{12}X$ compounds by Möller *et al.*[11] (RE is either Ce or Nd and A is one among Mn, Fe, Co, Ni, Ti, V, Cr, Cu, Zn, Al, Si, or P, and X is either B, C, or N). Giaremis *et al.*,[14] on the other hand, used an artificial neural network (ANN) to capture the relation between the structure of the Sm–Co permanent magnet and its saturation magnetization, and Wang *et al.*[15] developed an ML-based approach for predicting the saturation magnetization in Fe-based soft magnets. Other works on using data-driven approaches for predicting M_s are for phases other than 14:2:1 magnetic phases.[16,17] However, minimal research has been performed on applying the ML to predict intrinsic properties such as density and Curie temperature T_c using the elemental composition as features for 14:2:1 magnetic phases.[9,18] The attempt to predict the saturation magnetization for the 14:2:1 magnetic phase from chemical compositional features using ML-based approaches is the scope of this paper.

II. METHODOLOGY

A. Materials database

The materials database consists of literature data on 198 magnetic 14:2:1 phases and corresponding saturation magnetization M_s values in [μ_B/f.u.] measured at room temperature reported in distinct literature sources such as Herbst[1] and Burzo.[5] These phases comprise a specific combination of the 28 elements shown in Fig. 1(b). The dataset's M_s values in terms of [μ_B/f.u.] range from 5 to 35, as shown in Fig. 1(a). The dataset also has some phases with reported M_s values differing from one literature to the other, and such phases are treated as duplicates or repeated phases. This difference in the reported values for the same phase could be due to multiple reasons, such as using different measuring equipment or varying experimental conditions. As a pre-processing step, the elemental composition (C) was extracted for each phase in the dataset and correlated with the corresponding M_s values. The extracted elemental composition is composed of continuous variables serving as the input feature to train the supervised ML regression model with an expected output of M_s. Among the 198 phases in the dataset, 28 unique elements are obtained; hence, the feature vector size for the ML model is 28 with 1 target value. For example, the $Fe_{14}Nd_2B_1$ phase has a feature vector with 14 elements of Fe, 2 elements of Nd, 1 element of B, and the remaining 25 elements (features) as zero.

Furthermore, as seen in Table I, the dataset comprises phases with different components ranging from three (ternary) to six (senary). The dataset of 198 magnetic 14:2:1 phases has 143 unique phases and 55 repeated phases. The training data have 103 unique and 43 repeated phases, which are further split into the train and validation sets using the K-fold (K = 5) cross-validation method to

08 February 2024 09:36:03

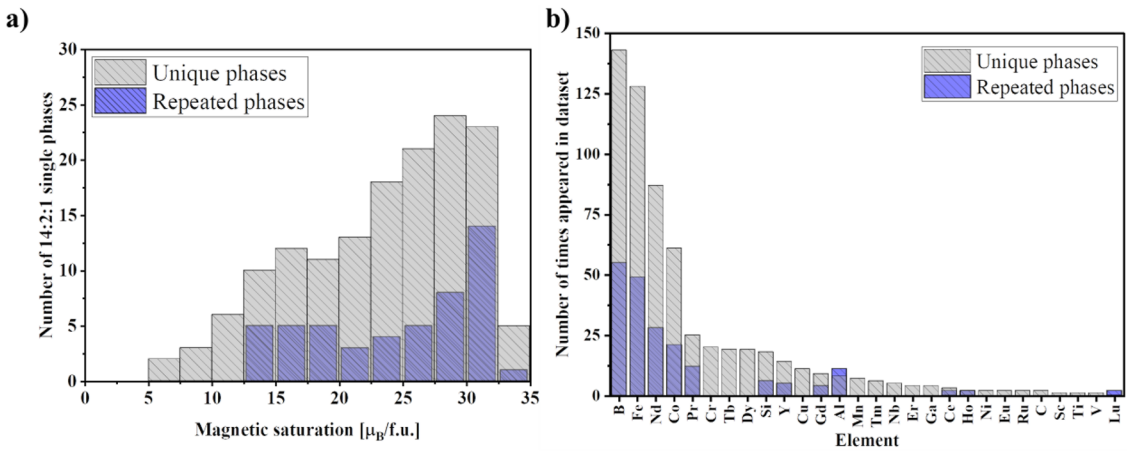

FIG. 1. Dataset containing 198 phases of 14:2:1 type. (a) Frequency distribution of saturation magnetization values in the dataset. (b) Frequency distribution of 28 elements observed across the dataset. The unique and repeated phases in the dataset are identified and highlighted in both distribution plots.

© Author(s) 2024

TABLE I. Number of phases of 14:2:1 type in the dataset categorized into different alloy systems. From three component system (ternary) to six component system (senary).

Alloy system	Number of phases in the dataset
Ternary	35
Quaternary	121
Quinary	40
Senary	2

evaluate the learning model's performance. The unseen test set has 52 phases with 40 unique phases to evaluate the performance of the trained model.

B. Machine learning models

Using the supervised learning approach, different ML regression models—namely the linear Huber regressor (HR),[19] tree-based approaches such as Random forest (RF),[20] Gradient boosting (GB),[21] Extreme gradient boosting (XGB),[22] Adaptive boosting with a decision tree as a base learner (AB-DT),[23] Multi-layer perceptron (MLP)[24] based on the artificial neural network, and weighted Voting regressor[25] (VR)—were trained and evaluated for performance on the validation set. Table II shows the optimized parameters of the different ML regressor models. Based on the performance of each model on the validation set, HR, RF, MLP, and AB-DT were considered base regressors for constructing a VR model. Note that the target property of magnetic saturation must be a continuous smooth function of substitutional chemical compositions. However, the inductive nature of base learners such as the tree-based regressors RF, GB, XGB, and AB-DT generates the step function, which is sometimes discontinuous,[26] while others (HR and MLP) tend to be a smooth continuous function. Nevertheless, VR combines the output from each base learner by taking the average, which leads to a smoother and continuous functional approximation when compared to individual base learner prediction.

Additionally, VR has the advantage that it can potentially overcome the drawbacks of individual base regressors to reduce the variance and leverage their strengths to strategically generalize well compared to stand-alone base learners. Though HR is robust to outliers and RF is good at capturing non-linear relationships, they have drawbacks in that they are prone to overfitting when intricate dependencies in data are not well captured by them.[19,20] MLP, with its complex transformation abilities, is sensitive to hyperparameter tuning. AB-DT's strength to enhance generalization ability by assigning more weight to data with higher errors has the drawback that it is sensitive to noise if decision trees are too complex.[23,24]

The performance evaluation of the regression model is possible using different metrics, of which mean absolute error (MAE), root mean squared error (RMSE), and correlation coefficient (R^2) are widely used metrics for comparing the regression model's performance. Each metric has its own importance and drawbacks; therefore, a combination of multiple metrics provides unbiased information on model performance,

$$MAE = \frac{1}{m}\sum_{j=1}^{m} |y_j - \hat{y}_j|$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{j=1}^{m} (y_j - \hat{y}_j)^2}$$

$$R^2 = 1 - \frac{\sum_{j=1}^{m} (y_j - p_k)^2}{\sum_{j=1}^{m} (\hat{y}_j - p_k)^2}$$

Here, m stands for the number of samples in the dataset, $\hat{y}_j$ is the jth target value, y_j is the predicted value for the jth sample, and p_k is the mean target value of all samples in the dataset.

The base regressors for the VR can either have a uniform or non-uniform weightage. The VR's final output is the average of the weighted predictions from the base regressors. If each base regressor has equal weightage, then it is a uniform VR; if weights vary, then it is a non-uniform weighted VR (NU-VR). As we have base regressors of different natures that include linear, tree-based, and ANN, using uniform VR would not be effective.[27,28] Figure 2 shows the working of the VR with four base regressors for M_s prediction from a given feature vector (f) extracted from the input chemical composition (C).

For a given input (C), each base regressor $F_i(f)$ predicts an output that gets multiplied by the assigned weight (w_i). VR calculates the average weighted prediction from each base regressor to obtain the final output M_s,

$$\text{Predicted } M_s = \frac{F_{HR}(f) * w_1 + F_{RF}(f) * w_2 + F_{MLP}(f) * w_3 + F_{AB-DT}(f) * w_4}{(w_1 + w_2 + w_3 + w_4)}$$
$$if\, w_1 = w_2 = w_3 = w_4 (\text{uniform } VR).$$

TABLE II. The optimized parameters of the different regressors used for the magnetic saturation predicting model. The optimized parameters are obtained after performance evaluation on the 5-fold cross-validation split.

Base regressor	Optimized parameters
HR	Epsilon: 1.35; max_iter: 200; alpha: 0.0001
RF	max_depth: 100; n_estimators: 100; criterion: mae; bootstrap: True
GB	n_estimator: 50; max_depth: 20; min_samples_leaf: 2; min_samples_split: 8
XGB	Objective: reg:squarederror; learning_rate: 0.05; max_depth: 100; alpha: 0.8; n_estimators: 100
AB-DT	Learning_rate: 0.01; n_estimators: 50; estimator: DecisionTreeRegressor(max_depth:10)
MLP	Activation: relu; hidden_layer_sizes: (50, 30); max_iter: 300; solver: lbfgs; tol: 0.0001; alpha: 0.0001

08 February 2024 09:36:03

© Author(s) 2024

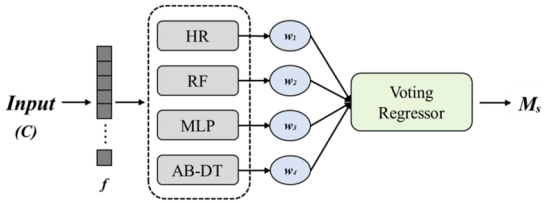

FIG. 2. Voting regressor pipeline for the M_s prediction model from the chemical composition. The average of the predicted M_s from the four base regressors with respective weights is considered the final output from the voting regressor.

There are different approaches for calculating the weightage, and it is mainly split into two categories: performance-based[27,29–32] and rank-based weights.[33,34] The performance-based approach considers the normalized performance values of each base learner as its weight. In a rank based approach, based on the cross-validation score, each base learner is assigned a rank. A better performing model is assigned the highest rank, implying more weightage. The approaches for weight calculation are common for both classification and regression tasks. The proposed approach considers both performance-based weights measured using the R^2 score and MAE and a rank-based weight measured using the RMSE on the validation set to obtain an optimized weight for each of the four shortlisted base regressors using the grid search techniques.

III. RESULTS

A. ML model training and cross-validation

Table III shows the performance of the trained regression models on the validation set using RMSE, MAE, and R^2 metrics. The base regressors HR, RF, MLP, and AB-DT considered for the voting regressor have validation errors of 1.08, 0.80, 0.71, and 0.75 [μ_B/f.u.] using MAE as metric and 1.17, 1.17, 1.09, and 1.08 [μ_B/f.u.] using the RMSE as metric, respectively. The other regression models, such as GB and XGB, have validation errors of 1.50 and 1.16 [μ_B/f.u.] using MAE and 1.90 and 1.61 [μ_B/f.u.] using RMSE as metrics, respectively. The voting regressor model using the base regressors was constructed with uniform and non-uniform weights. For uniform weighted VR, all four base regressors were assigned an equal weightage of 1; for non-uniform weighted VR, base regressors HR, RF, MLP, and AB-DT were assigned weightage of 0.89, 0.94, 0.95, and 0.96, respectively.

TABLE III. Performance comparison between the different base regressor models and weighted regressor models in terms of coefficient of fit (R^2), root mean squared error (RMSE), and mean absolute error (MAE) on the validation dataset having 29 phases.

Regression models	R^2	RMSE	MAE
Huber regressor (HR)	0.97	1.17	1.08
Random forest (RF)	0.97	1.17	0.80
Gradient boosting (GB)	0.93	1.90	1.50
Extreme gradient boosting (XGB)	0.91	1.61	1.16
Multi-layer perceptron (MLP)	0.96	1.09	0.71
Decision tree + adaptive boosting (AB-DT)	0.98	1.08	0.75
VR-non-uniform weight (NU-VR)	0.97	1.07	0.72
VR-uniform weight (U-VR)	0.94	1.16	0.79

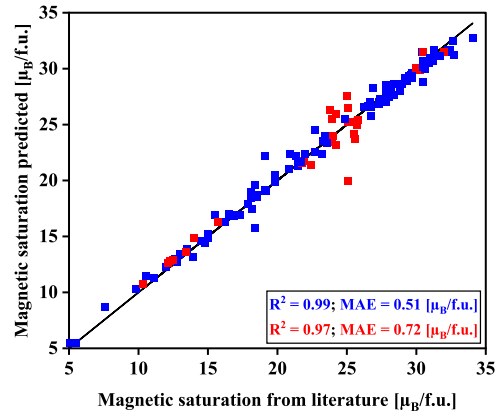

FIG. 3. Non-uniform weighted voting regressor model fitting of the saturation magnetization on the training set (in blue color) containing 117 phases in terms of mean absolute error (MAE) and R^2 score and on an independent validation subset (in red color) of 29 phases from one of the 5 folds.

The uniform weighted VR has a validation error of 0.79 and 1.16 [μ_B/f.u.] using MAE and RMSE, respectively. For the same validation set, the non-uniform weight VR resulted in an error of 0.72 and 1.07 [μ_B/f.u.] using MAE and RMSE, respectively. The validation error using the uniform weighted VR resulted in an average deviation of 3%, whereas the non-uniform weighted VR is less than 3%. Due to the superior performance of non-uniform weighted VR (NU-VR), it has been considered for further testing. Figure 3 shows the performance of the trained model NU-VR on a training set of 117 phases combined with the validation set of 29 phases against the reference values from different literature sources for one of the 5 folds. A low overall MAE of 0.72 [μ_B/f.u.] and a good fit of 0.97 using the R^2 score were measured for the training and validation sets. The model fits the dataset well and results in an absolute error of less than 2 [μ_B/f.u.] for most phases and a relatively higher error of more than 2 [μ_B/f.u.] for 3 phases out of 143 phases. The phases $Fe_{10}CoAl_3Nd_2B$, $Fe_{14}Nd_{1.6}Er_{0.4}B$, and $Fe_{14}Nd_{1.5}Tb_{0.5}B$ have an absolute error of 5.16, 2.6, and 2.49 [μ_B/f.u.], respectively.

B. ML model testing on unseen test compositions

On the unseen test set, the trained model resulted in an error of 0.77 and 0.88 [μ_B/f.u.] using MAE and RMSE, respectively. The resulting error would mean the average deviation from the trained model is 3%. The low error on the unseen test set confirms that the model has generalized well to the dataset. Furthermore, the model's performance on the repeated and unique phases in the unseen test set was calculated. Figure 4 shows the performance of the trained model NU-VR on the unseen test set. For the 40 unique phases, an MAE of 0.8 and an RMSE of 0.9 [μ_B/f.u.] have been obtained. The model's performance on the 12 repeated phases among the 52 phases in the unseen test resulted in a low MAE of 0.7 and RMSE of 0.8 [μ_B/f.u.]. Note that the repeated 12 phases in the

08 February 2024 09:36:03

© Author(s) 2024

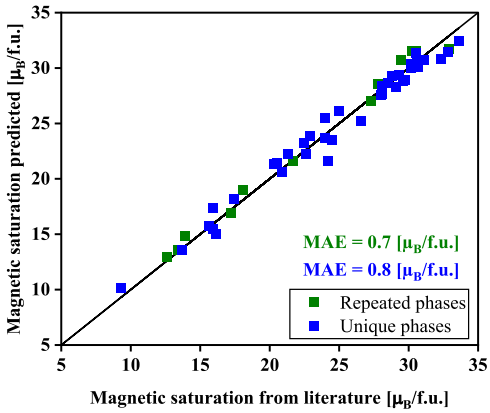

FIG. 4. Non-uniform weighted voting regressor model predictions on an unseen test set (52 phases). The test set contains 39 unique and 13 repeated phases taken from independent literature sources that were not part of the training set.

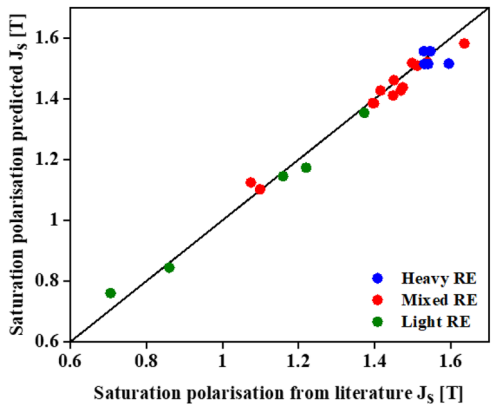

FIG. 5. Compares the calculated saturation polarization J_s [T] predicted from the trained machine learning models and literature values. The phases are categorized into 14:2:1 phases containing heavy rare earth (in blue color), light rare earth (in green color), and mixed (heavy and light combined) rare earth metals (in red color). The calculated J_s in [T] is obtained using the saturation magnetization M_s in [μ_B/f.u.] and density ρ in [g/cm^3].

unseen test set are obtained from the different literature sources that have reported slightly different M_s values. Only one unique phase has an absolute error of more than 2 [μ_B/f.u.] and has been identified as $Fe_{12}Cu_2Y_2B$. The remaining 51 phases have an absolute error of less than 2 [μ_B/f.u.], with most having an error of less than 1 [μ_B/f.u.].

C. Application of the trained model to compare the measured and predicted saturation polarization

Magnetic saturation polarization is often calculated from the measured or experimental saturation magnetization and density using the physical formula, depending on the units of the measured M_s. Most works in the literature report the measured M_s value in [μ_B/f.u.] or emu, which requires the density in [g/cm^3] to calculate the J_s [T] using the physical formula provided in the literature,[1,5]

$$J_s = \frac{m * \rho * 4\pi}{10\,000}$$

$$m = \frac{M_s * 9.274 * 10^{-24}}{(\sum A) * 1.6605 * 10^{-27}}$$

Here, m refers to the saturation magnetization in emu/g, ρ is density in [g/cm^3], and M_s is saturation magnetization in [μ_B/f.u.]. Moreover, A is the formula mass of the 14:2:1 phase.

J_s can also be calculated from M_s using another physical formula reported in the literature,[1,5] provided M_s has been measured or reported in Ampere per meter [A/m],

$$J_s = \mu_0 * M_s$$

where M_s is saturation magnetization measured in [A/m] and μ_0 is vacuum permeability in [V$_s$/Am].

Using the developed model to predict the M_s, J_s can be calculated for the phases whose density ρ has been measured or reported in the literature. Similarly, using the approach suggested by Kini *et al.*[18] for the density prediction model using chemical composition, the density can be predicted for the alloys whose M_s in [μ_B/f.u.] has already been reported in the literature for calculating the J_s. Note that the developed model for M_s prediction and the existing density prediction model from Kini *et al.*[18] have been trained on different chemical compositions that resulted in a different chemical elemental feature vector. Therefore, to predict M_s, or the density of the unknown phase, the input feature, which is the chemical composition, should be within the feature space of the chemical compositions that were used for training the models.

Out of the 143 unique phases in the dataset for the M_s prediction model, there are 24 unique phases whose density and M_s values have both been reported in the literature. These 24 phases belong to ternary, quaternary, and quinary alloy systems. Note that for these 24 phases, the literature sources for reported M_s and density are different, implying that the calculated J_s values may have some degree of error due to differences in experimental setup or subjective error. Ideally, the J_s value calculated using the M_s and density from the literature source should match the J_s value calculated from the predicted M_s and density values from the respective trained ML model.

Comparing the resulting J_s from the literature and predicting them would serve as a physics-based or human-in-loop sanity check for the dataset used for training the ML models. Figure 5 compares the calculated J_s in [T] using M_s in [μ_B/f.u.] and density in [g/cm^3] from literature sources and the trained model. Depending on the presence of the rare earth (RE) element, 24 phases are categorized into heavy RE, light RE, and mixed RE. An average absolute deviation of 2%, 2%, and 3% was seen among the phases with light RE, mixed RE, and heavy RE. A maximum deviation of 0.1 T was seen among the 24 phases tested.

IV. DISCUSSION

The dataset for the M_s prediction model is diverse, with phases belonging to different alloy systems and phases with different rare

08 February 2024 09:36:03

© Author(s) 2024

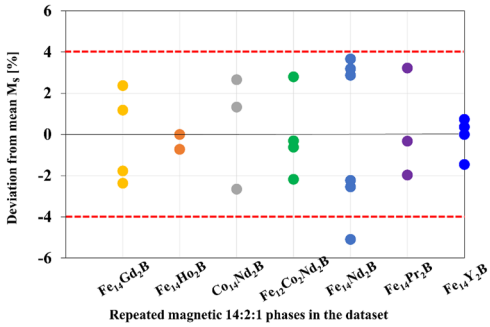

FIG. 6. Example of the reported magnetic saturation [μ_B/f.u.] at room temperature for the repeated magnetic 14:2:1 phases in the dataset collected from the various literature sources. For each repeated phase, the deviation of phases from their mean value is calculated, and a maximum deviation of 5% is observed.

earth elements. Furthermore, the repeated phases reported different M_s values in [μ_B/f.u.] for the same chemical composition, and it has been primarily due to the different measurement methods and equipment, errors in measurement, and considerable deviation in the ambient temperature for conducting experiments. In the literature sources, the reported M_s values are measured using saturation law or with a vibrating sample magnetometer (VSM) or extraction magnetometer at different external fields such as 1.0–8.0 T and varying ambient temperatures between 295 and 305 K.

Figure 6 shows the magnetic saturation in [μ_B/f.u.] for the repeated magnetic 14:2:1 phases in the training and test sets for seven different phases. There are some other repeated phases in the dataset apart from the seven phases shown in Fig. 6, but for explanation, we have chosen phases that are repeated in varying numbers. The phases $Fe_{14}Ho_2B$ appear twice, $Fe_{14}Gd_2B$, $Fe_{14}Pr_2B$, and $Co_{14}Nd_2B$ four times, $Fe_{14}Y_2B$ five times, and $Fe_{14}Nd_2B$ six times in the dataset. The highest standard deviation among repeated phases is 2 [μ_B/f.u.] for $Fe_{14}Nd_2B$, and the least is 0.1 [μ_B/f.u.] for $Fe_{14}Ho_2B$. Since these phases are from different literature sources, they have the same chemical composition (input features) but different M_s values. The mean value of repeated phases is calculated, and the deviation of each repeated phase from its mean value is noted. The maximum deviation of 5% was for the $Fe_{14}Nd_2B$. For other repeated phases, the deviation is less than 5%. If the range or spread of the reported M_s value for a given phase is high, resulting in clusters, then the model learns the optimal M_s value from the cluster with more repeated phases using a majority vote. However, suppose there is no possibility of forming clusters for reported M_s values for a given phase. In that case, the model learns the optimal M_s value such that it has the least error concerning most of the repeated phases. Moreover, the iterative learning characteristic of most of the ML models ensures that the presence of outliers in repeated phases does not influence the decision-making process of the ML model.

A. Generalization ability of the trained model

The base regressors considered for constructing the voting regressor model include parametric and non-parametric models. Among the base regressors, RF and AB-DT are tree-based models, HR is a linear model, and MLP is an artificial neural network capable of learning non-linear functions. With such diverse base regressors, assigning equal weightage to each one is difficult as they differ greatly in the learning process. Therefore, unequal weights have been assigned to obtain a smoother and continuous function to have a generalized and robust model that has proven to fit the dataset better than using a uniform VR.

The low MAE and RMSE for the phases in the validation dataset suggest that the model did well in generalizing the diverse dataset with different alloy systems. However, despite the low prediction error of the trained non-uniform weighting VR model, there has been a noticeable difference in the MAE and RMSE values. RMSE penalizes the phases with high error and suggests the presence of outliers or phases on which the model did not fit well. A relatively higher RMSE of 1.07 [μ_B/f.u.] than MAE of 0.72 [μ_B/f.u.] on the validation set is due to the presence of three phases, $Fe_{10}CoAl_3Nd_2B$, $Fe_{14}Nd_{1.6}Er_{0.4}B$, and $Fe_{14}Nd_{1.5}Tb_{0.5}B$, that have an absolute error of more than 2 [μ_B/f.u.], of which $Fe_{10}CoAl_3Nd_2B$ resulted in an absolute error of 5.16 [μ_B/f.u.]. In the training set, the phases where the iron (Fe) element is substituted by different percentages of Co and Al elements occur only three times.

Figure 7 shows that a higher amount of aluminum (Al) in the phase has a negative impact and could lead to a low M_s value. The Al substitutes 3 out of 4 elements of Fe in $Fe_{10}CoAl_3Nd_2B$. It has a higher negative impact on the prediction, leading to a lower M_s value of 19.94 [μ_B/f.u.] as compared to the ground truth of 25.1 [μ_B/f.u.]. However, including more such phases in the training set could assist the learning model in generalizing better. The presence of such phases in lower numbers in a dataset is challenging as they lie on the solubility limit's boundary and require more training phases to capture the data fit.

Figure 7 shows the feature importance plot calculated using the SHapley Additive exPlanations (SHAP)[35] method, with the SHAP value indicating the importance of the individual element. Since the base learners for the VR model have linear, non-linear, and tree-based models, a model-agnostic approach is recommended to get

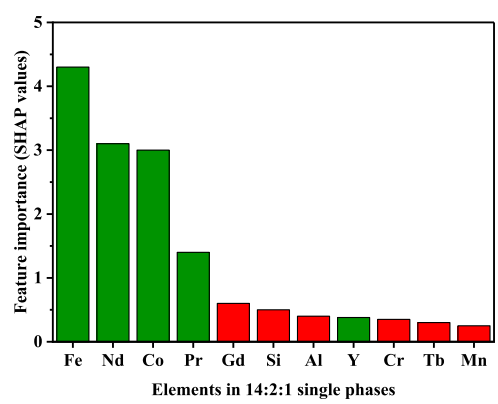

FIG. 7. Global feature importance plot for the trained weighted voting regressor obtained using the SHAP method. The feature importance values are mean absolute SHAP values; green indicates a positive impact, and red suggests a negative impact on the M_s prediction.

08 February 2024 09:36:03

© Author(s) 2024

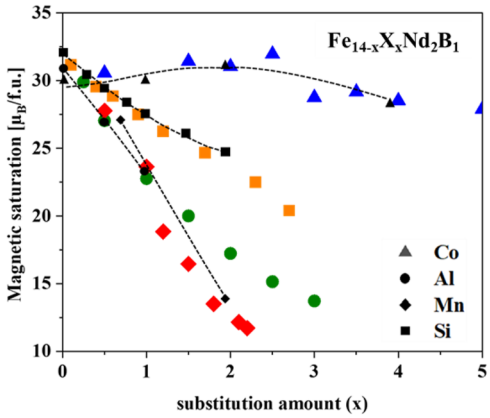

FIG. 8. Comparison between the predicted (colored data points) and reported (black points) M_s from an unseen source for the $Fe_{14-x}X_xNd_2B_1$ (X = additional elements). The trained model has predicted the magnetic saturation for the higher substitution of elements (Co, Al, Mn, and Si) within the solubility limits.

feature importance. Note that the plot shows only the top 11 out of 28 features used for model development and explains the trained model's behavior but not data distribution. The color green indicates a positive impact, whereas red indicates a negative impact on the model's prediction. For example, if a phase has Nd, Fe, and Si, then Nd and Fe would cause the model to predict a high M_s value, whereas Si would tend to lower the M_s value prediction, and this effect is proportional to the percentage of each element in the phase. This also agrees well with the experimental results shown in Fig. 8 from Burzo *et al.*[5]

The trained model performance on the unseen test set on unique and repeated phases suggests that the model fits well on a dataset, showing low MAE and RMSE errors. There has been no significant difference in the MAE and RMSE values for phases in the test set. Furthermore, the low MAE of 0.8 [μ_B/f.u.] on unique phases indicates that the model is not overfitting and can capture the relationship between input features and the M_s value. The only phase that resulted in a higher error is $Fe_{12}Cu_2Y_2B$, and it has been found that this composition is not a single phase due to the limited solubility of Cu.[4]

Figure 8 compares the literature reported[1,5] and predicted M_s values for different substitution elements, which include Co, Al, Mn, and Si. The literature reported values marked in black and predicted in different colors. The metals Mn and Al have high slopes, suggesting greater sensitivity than Co or Si.

Furthermore, an attempt has been made to extend the existing plot from Burzo[5] by predicting M_s values for higher substitution amounts and within the solubility limit. The dotted lines represent the behavior of $Fe_{14-x}X_xNd_2B_1$ for increasing substitution amounts of X elements (Co, Al, Mn, and Si) reported by Burzo.[5]

B. Effect of aggregation of data based on the chemical composition

The presence of unique and repeated phases in the dataset primarily provides the model with data that is closely representative of real-world data and includes variability in measurements from different sources. Furthermore, aggregating the data using a single metric such as mean or average would lead to a true M_s value for some compositions, and for others, this may lead to a high error.[9] Figure 6 shows that the deviation error from the mean or average (metric) value within each repeated phase varies non-uniformly. This could imply that an alternative metric such as median, mode, or any custom metric is needed to aggregate the composition where the metric of mean is not suitable. Suggesting that it is not recommended to have a single metric for aggregating all the repeated compositions in the dataset. Note that aggregation based on a single metric would also induce bias and overfit the model with high variance. It is well known that the measure of the trained model's ability to capture the universal representation of the dataset is based on the predictability of the unseen test set,[36,37] and the low MAE of 0.88 [μ_B/f.u.] on unique and 0.78 [μ_B/f.u.] on repeated phases in the test set confirms that there is no bias toward repeated phases and the model generalized well. This highlights the importance of the VR model, which strategically handles the challenges of dataset characteristics by leveraging the strengths of base learners.

Moreover, with the non-aggregated dataset, the learning model can capture the composition dependence on an error, and its effect is being evaluated using the MAE and RMSE. Finally, it must be noted that the supervised ML model's learning process is such that for the duplicate or repeated data points, it learns the optimized target value, which is close to the cluster of duplicate target values, by minimizing the error during the training phase. Additionally, in tree-based base regressor models such as RF and AB-DT in VR, the effect due to the presence of outliers or noise in the repeated phases is reduced as they are treated internally.

V. STRENGTHS AND LIMITATIONS OF THE CURRENT MODEL

The dataset used for developing the model has some compositions with multiple entries (repeated phases) as they have varying M_s values from different literature sources. The variation in the reported values for the same composition from such literature sources has a standard deviation of 2 [μ_B/f.u.]. Nevertheless, the trained model could predict values on unseen unique phases with a low MAE of 0.8 [μ_B/f.u.].

The developed model has been trained on a feature set of 28 elements shown in Fig. 1(b) and compositions within the solubility's boundary limits. Therefore, we have noted that phases with a composition that lies close to the solubility limit or even outside produce a comparatively high error. Adding more training samples for such compositions can significantly improve the model's prediction capability.

VI. CONCLUSIONS

1. The model for predicting the M_s values for the magnetic 14:2:1 phases has been developed using the supervised learning method with chemical composition as input features, resulting in a low MAE of 0.8 [μ_B/f.u.] on the unseen test set, which accounts for the average deviation of 3% when compared to the reference values from literature sources.
2. The trained model is a single generalized model for magnetic 14:2:1 phases applicable for ternary to senary alloys having

08 February 2024 09:36:03

© Author(s) 2024

heavy and light RE that works within the solubility limit of the 28 elements used for training the model. Furthermore, the trained model has generalized well on the dataset and does not get affected by the presence of different heavy and light RE substituted 14:2:1 phases.

3. The proposed approach of combining parametric and non-parametric models as base regressors for building a non-uniform weighted voting regressor model has proven more effective on diverse datasets involving unique and repeated phases. Out of 198 phases, only one phase, i.e., $Fe_{10}CoAl_3Nd_2B$, has resulted in a higher absolute error of more than 5 [μ_B/f.u.].
4. The comparison between the J_s [T] calculated from the measured M_s and density values from the literature source and the predicted values resulted in a MAE of 0.1 T. Such a comparison could serve as the basis for developing a physics-informed ML model or a tool for a sanity check of the dataset used for training the models, as a higher difference indicates the possible presence of measuring errors.

ACKNOWLEDGMENTS

This work was performed within the scope of the "MEMORI" project. The "MEMORI" project was made possible by funding from the Carl Zeiss Foundation (Grant No. P2018-03-002). Additionally, this publication was funded by Aalen University of Applied Sciences and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - (Project no: 512645013). The authors would like to acknowledge Anoop Kini for providing valuable feedback on the results.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Amit Kumar Choudhary: Conceptualization (equal); Data curation (equal); Formal analysis (lead); Investigation (equal); Methodology (lead); Software (lead); Validation (equal); Visualization (equal); Writing – original draft (lead); Writing – review & editing (equal). **Dominic Hohs**: Data curation (equal); Resources (supporting); Validation (equal); Writing – review & editing (supporting). **Andreas Jansche**: Formal analysis (supporting); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Validation (supporting). **Timo Bernthaler**: Conceptualization (supporting); Funding acquisition (equal); Project administration (equal); Supervision (equal); Validation (supporting); Writing – review & editing (supporting). **Dagmar Goll**: Conceptualization (equal); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Methodology (equal); Project administration (lead); Resources (lead); Supervision (equal); Validation (equal); Writing – review & editing (equal). **Gerhard Schneider**: Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Project administration (equal); Resources (equal); Supervision (equal); Validation (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are not publicly available due to privacy or ethical reasons, and the data forms part of an ongoing study.

REFERENCES

[1] J. F. Herbst, "$R_2Fe_{14}B$ materials: Intrinsic properties and technological aspects," Rev. Mod. Phys. **63**(4), 819–898 (1991).
[2] C. Abache and H. Oesterreicher, "Structural and magnetic properties of $R_2Fe_{14-x}T_xB$ (R = Nd, Y; T = Cr, Mn, Co, Ni, Al)," J. Appl. Phys. **60**(3), 1114–1117 (1986).
[3] A. Kowalczyk and A. Wrzeciono, "Magnetic and crystallographic properties of $R_2Fe_{14-x}B$ compounds (R = Y, Nd, and Gd)," Phys. Status Solidi A **110**(1), 241–245 (1988).
[4] A. Kowalczyk and A. Wrzeciono, "Structural and magnetic characteristics of $R_2Fe_{14-x}Cu_xB$ systems (R = Y, Nd and Gd)," J. Magn. Magn. Mater. **74**(3), 260–262 (1988).
[5] E. Burzo, "Permanent magnets based on R-Fe-B and R-Fe-C alloys," Rep. Prog. Phys. **61**(9), 1099 (1998).
[6] A. Kowalczyk, "Magnetic and crystallographic properties of substituted $Pr_2Fe_{14-x}M_xB$ compounds (M = Si, Ga, Cr and Cu)," J. Magn. Magn. Mater. **82**(1), L1–L4 (1989).
[7] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, "The high-throughput highway to computational materials design," Nat. Mater. **12**(3), 191–201 (2013).
[8] J. Nelson and S. Sanvito, "Predicting the Curie temperature of ferromagnets using machine learning," Phys. Rev. Mater. **3**(10), 104405 (2019).
[9] A. K. Choudhary, A. Kini, D. Hohs, A. Jansche, T. Bernthaler, O. Csiszár, D. Goll, and G. Schneider, "Machine learning-based Curie temperature prediction for magnetic 14:2:1 phases," AIP Adv. **13**(3), 035112 (2023).
[10] A. Agrawal, P. D. Deshpande, A. Cecen, G. P. Basavarsu, A. N. Choudhary, and S. R. Kalidindi, "Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters," Integr. Mater. Manuf. Innovations **3**(1), 90–108 (2014).
[11] J. J. Möller, W. Körner, G. Krugel, D. F. Urban, and C. Elsässer, "Compositional optimization of hard-magnetic phases with machine-learning models," Acta Mater. **153**, 53–61 (2018).
[12] Y. Zhang and X. Xu, "Predicting magnetic remanence of NdFeB magnets from composition," J. Supercond. Novel Magn. **34**(11), 2711–2715 (2021).
[13] A. Vishina, O. Y. Vekilova, T. Björkman, A. Bergman, H. C. Herper, and O. Eriksson, "High-throughput and data-mining approach to predict new rare-earth free permanent magnets," Phys. Rev. B **101**(9), 094407 (2020).
[14] S. Giaremis, G. Katsikas, G. Sempros, M. Gjoka, C. Sarafidis, and J. Kioseoglou, "*Ab initio*, artificial neural network predictions and experimental synthesis of mischmetal alloying in Sm–Co permanent magnets," Nanoscale **14**(15), 5824–5839 (2022).
[15] Y. Wang, Y. Tian, T. Kirk, O. Laris, J. H. Ross, R. D. Noebe, V. Keylin, and R. Arróyave, "Accelerated design of Fe-based soft magnetic materials using machine learning and stochastic optimization," Acta Mater. **194**, 144–155 (2020).
[16] T. D. Rhone, W. Chen, S. Desai, S. B. Torrisi, D. T. Larson, A. Yacoby, and E. Kaxiras, "Data-driven studies of magnetic two-dimensional materials," Sci. Rep. **10**(1), 15795 (2020).
[17] S.-O. Kaba, B. Groleau-Paré, M.-A. Gauthier, A.-M. Tremblay, S. Verret, and C. Gauvin-Ndiaye, "Prediction of large magnetic moment materials with graph neural networks and random forests," Phys. Rev. Mater. **7**, 044407 (2022).

08 February 2024 09:36:03

© Author(s) 2024

[18]A. Kini, A. Kumar Choudhary, D. Hohs, A. Jansche, H. Baumgartl, R. Büttner, T. Bernthaler, D. Goll, and G. Schneider, "Machine learning-based mass density model for hard magnetic 14:2:1 phases using chemical composition-based features," Chem. Phys. Lett. **811**, 140231 (2023).
[19]A. B. Owen, in *Contemporary Mathematics*, edited by J. S. Verducci, X. Shen, and J. Lafferty (American Mathematical Society, Providence, RI, 2007), pp. 59–71.
[20]L. Breiman, "Random forests," Mach. Learn. **45**(1), 5–32 (2001).
[21]J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Ann. Stat. **29**(5), 1189–1232 (2001).
[22]C. Cervellera and D. Macciò, "Gradient boosting with extreme learning machines for the optimization of nonlinear functionals," in *Advances in Optimization and Decision Science for Society, Services and Enterprises*, 3 (Springer, Cham, 2019).
[23]R. Freund, P. Grigas, and R. Mazumder, "Adaboost and forward stagewise regression are first-order convex optimization methods," arXiv:1307.1192 (2013).
[24]S. K. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification," IEEE Trans. Neural Networks **3**(5), 683–697 (1992).
[25]K. An and J. Meng, in *Advanced Intelligent Computing Theories and Applications*, edited by D.-S. Huang, Z. Zhao, V. Bevilacqua, and J. C. Figueroa (Springer, Berlin, Heidelberg, 2010), pp. 540–546.
[26]J. M. Chambers and T. J. Hastie, *Statistical Models in S* (Routledge, Boca Raton, 2017).
[27]M. Shahhosseini, G. Hu, and H. Pham, "Optimizing ensemble weights and hyperparameters of machine learning models for regression problems," Mach. Learn. Appl. **7**, 100251 (2022).
[28]M. Shahhosseini, R. A. Martinez-Feria, G. Hu, and S. V. Archontoulis, "Maize yield and nitrate loss prediction with machine learning algorithms," Environ. Res. Lett. **14**(12), 124026 (2019).
[29]V. C. Osamor and A. F. Okezie, "Enhancing the weighted voting ensemble algorithm for tuberculosis predictive diagnosis," Sci. Rep. **11**(1), 14806 (2021).
[30]C. Shikun and N. M. Luc, "RRMSE voting regressor: A weighting function based improvement to ensemble regression," arXiv:2207.04837 (2022).
[31]V. Nourani, G. Elkiran, and S. I. Abba, "Wastewater treatment plant performance analysis using artificial intelligence—An ensemble approach," Water Sci. Technol. **78**(10), 2064–2076 (2018).
[32]Y. Li and Y. Luo, "Performance-weighted-voting model: An ensemble machine learning method for cancer type classification using whole-exome sequencing mutation," Quant. Biol. **8**(4), 347–358 (2020).
[33]B. Erdebilli and B. Devrim-İçtenbaş, "Ensemble voting regression based on machine learning for predicting medical waste: A case from Turkey," Mathematics **10**(14), 2466 (2022).
[34]Y. Zhao, Z. Shi, J. Zhang, D. Chen, and L. Gu, "A novel active learning framework for classification: Using weighted rank aggregation to achieve multiple query criteria," Pattern Recognit. **93**, 581–602 (2019).
[35]S. M. Lundberg and S.-I. Lee, *Advances in Neural Information Processing Systems* (Curran Associates, Inc., 2017).
[36]Y. V. Karpievitch, E. G. Hill, A. P. Leclerc, A. R. Dabney, and J. S. Almeida, "An introspective comparison of random forest-based classifiers for the analysis of cluster-correlated data by way of RF++," PLoS One **4**(9), e7087 (2009).
[37]D. H. Wolpert, "Stacked generalization," Neural Networks **5**(2), 241–259 (1992).

08 February 2024 09:36:03

© Author(s) 2024