Big Data and Information Analytics

BDIA, 9: 285–301.
DOI: 10.3934/bdia.2025013
Received: 12 November 2025
Revised: 20 November 2025
Accepted: 23 November 2025
Published: 01 December 2025

https://www.aimspress.com/journal/BDIA

Research article

Multi-objective optimization design of process parameters for rare earth hard magnetic materials based on machine learning

Haozhen Hu, Mengxian Zhao and Cun Chen*

School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, 450001, China

* **Correspondence:** Email: chencun@zzu.edu.cn.

Abstract: In the present work, combining machine learning and optimization theory, a novel framework is designed to optimize the process parameters of rare earth hard magnetic materials $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$ with multi-objective performance. Considering the small dataset problem, the interpolation and extrapolation technique was introduced for data augmentation. Based on the expanded dataset, predictive models for coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) concerning heat treatment parameters (annealing temperature and magnetic field strength) are established. Based on the enhanced data, the prediction accuracy of the model for H_{cj} and $(BH)_{max}$ reaches 0.980 and 0.997, respectively. After that, the NSGA-II algorithm is utilized for multi-objective optimization; the top 12 optimal process parameter combinations are recommended with higher coercivity and maximum magnetic energy product. This study provides a new way for the performance prediction and process optimization of rare earth hard magnetic materials, which can accelerate the design of materials with multi-objective performance.

Keywords: machine learning, rare-earth hard magnetic materials, coercivity, maximum magnetic energy product, multi-objective optimization, data augmentation

1. Introduction

Neodymium-iron-boron (NdFeB) permanent magnet materials are important rare earth functional materials, which can be used in automation, electrical appliances, robotics, medical and industrial equipment, as well as electric motors, due to the excellent magnetic field strength and coercivity [1]. Compared with the traditional asynchronous motor, rare earth permanent magnet synchronous motor

has the advantages of high efficiency, high power factor, high power density, and high reliability and has a better energy-saving effect. NdFeB permanent magnets demonstrate great potential for various applications owing to their excellent comprehensive properties. As the third-generation magnetic material, NdFeB permanent magnets possess several advantages and are widely utilized due to their relatively low production costs. To meet the higher requirements of applications, it is necessary to design NdFeB permanent magnets with higher coercivity (H_{cj}) and maximum magnetic energy product ((BH)$_{max}$). This will contribute to achieving breakthroughs in the field of applications. To improve the overall performance of NdFeB permanent magnets, how to design the parameters to achieve multi-objective optimization is a challenging issue.

In addition to the influence of composition on magnet properties [2–4], the process parameters also have a significant impact on magnet properties [5–12]. The traditional design of NdFeB magnets tends to rely on intuition or the trial-and-error method [13,14]. However, the space of possible process parameter compositions is huge, making traditional methods inadequate. Intuitive methods may lead to errors in the prediction of significant components and reduce reliability. Meanwhile, the traditional trial-and-error method is time-consuming and inefficient. Therefore, the development of an efficient method for designing high-performance and low-cost rare earth permanent magnet materials is of great importance. With the development of data science, machine learning has opened up a new blueprint for discovering and rationally designing high-performance materials. Machine learning models can quickly reveal the implicit relationships between composition, process parameters, and properties in the existing datasets to accurately predict material properties, thereby accelerating the design and development of new materials. The application of machine learning in materials science spans diverse material systems, such as amorphous materials [15], steels [16–21], and high entropy alloys [22–27]. The successful application of machine learning as a data analysis technology in material development provides insight into the optimization of process parameters for magnetic materials.

Gu et al. [28] studied the effect of magnetic field intensity on the distribution, structure, and hard magnetism of a magnets experimentally and found that magnetic heat treatment can not only contribute to the formation of uniformly distributed B-rich zone but also benefit the precipitation of α-Fe. An et al. [29] observed the interaction domain structure of melt spinning and annealing samples at 673 K and confirmed the existence of exchange coupling interaction. The exchange coupling effect is enhanced by grain refinement of magnetic annealed samples. The residual magnetism and the maximum magnetic energy product increase by 4% and 6%, respectively. Scholars are trying to use machine learning to study magnetic materials. Wang et al. [30] applied machine learning models to predict magnetic saturation (B_S), coercive (magnetic) force (H_C), and magnetostriction (λ), and used a stochastic optimization framework to optimize the corresponding magnetic properties. Lambardet et al. [31] used machine learning and Bayesian optimization to optimize the direct extrusion process of NdFeB magnets. At present, the research on coercivity and maximum magnetic energy product of rare earth hard magnetic alloys is still limited to a single target performance optimization. The multi-objective performance optimization of NdFeB magnetic materials with respect to process parameters cannot be achieved by the traditional trial-and-error method. How to optimize the multi-objective performance of NdFeB magnets is still an unsolved problem.

In this paper, machine learning and multi-objective optimization algorithm are combined to achieve the performance optimization of coercivity and maximum magnetic energy product for NdFeB permanent magnets. This study focuses on the optimization of heat treatment parameters of $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$ to improve both coercivity and maximum magnetic energy product. The heat

treatment parameters, especially the effects of temperature and magnetic field strength on the properties of the materials, are studied. Due to it is time-consuming and costly to obtain a large amount of experimental data containing process parameters, the available data for process parameters-to-performance relationships are limited. To solve the problem of small data sets, the data are enhanced by cubic spline interpolation and polynomial extrapolation. The machine learning model is established, in which XGBoost has the highest prediction accuracy, and the coercive force and maximum magnetic energy product R^2 are 0.980 and 0.997, respectively. The multi-objective optimization algorithm NSGA-II is employed, and a set of 12 recommended heat treatment parameters is provided. The coercive force and maximum magnetic energy product of the optimized magnet can be increased by 4.7% and 7.4%, respectively. In summary, this study improves the coercive force and maximum magnetic energy product systematically using a multi-objective optimization method for NdFeB permanent magnets.

2. Design strategy and methods

Figure 1 presents a design framework that utilizes heat treatment parameters to achieve simultaneous optimization of coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) for rare-earth hard magnetic material $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$. The framework includes four main modules: database establishment, data augmentation, model building, and multi-objective optimization of heat treatment parameters. Finally, the process parameters of the material are recommended, which can achieve better comprehensive performance.

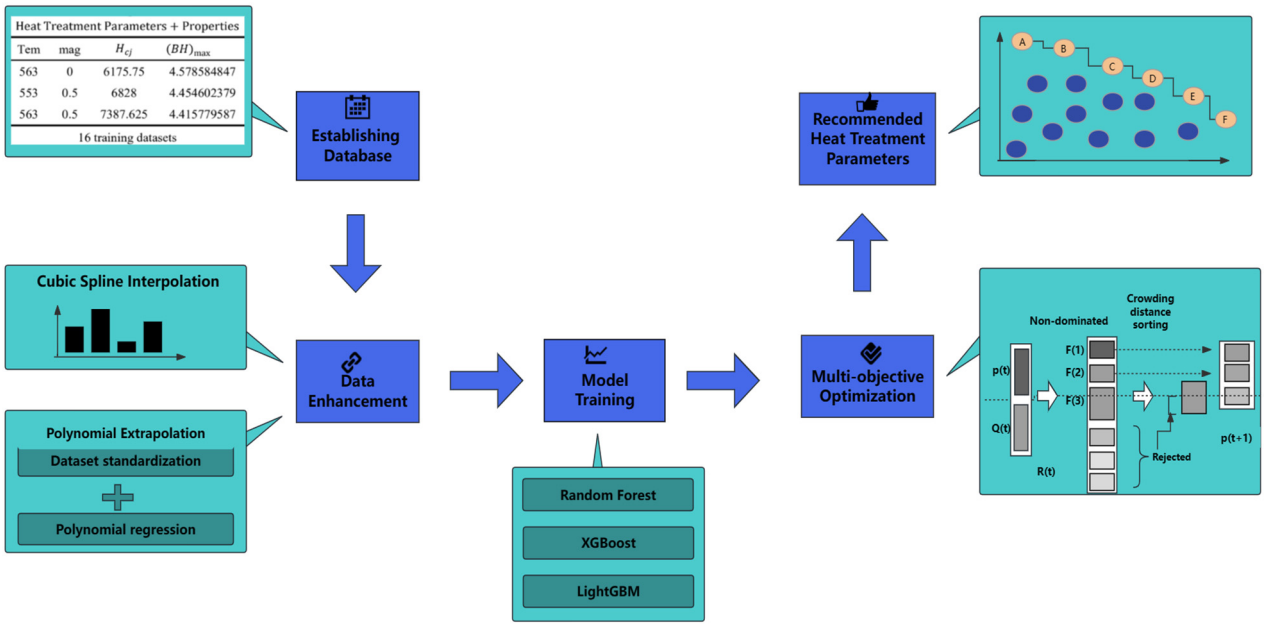

Figure 1. Frame diagram of the multi-objective optimization design.

2.1. Data collection

We collect 16 data points related to heat treatment parameters (annealing temperature and magnetic field intensity) and performance from reference [28], which is shown in Table 1.

Table 1. Heat treatment parameters and corresponding performance for 16 samples of $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$.

Temperatures(K)	magnetic field(T)	$H_{cj}(Oe)$	$(BH)_{max}(MGOe)$
563	0	6175.750	4.579
553	0.5	6828.000	4.455
563	0.5	7387.625	4.416
573	0.5	6214.875	4.430
583	0.5	8552.750	4.352
593	0.5	6434.000	4.562
553	0.8	8324.000	4.778
563	0.8	7819.375	4.447
573	0.8	6720.750	5.014
583	0.8	7605.500	4.823
593	0.8	7018.625	4.566
553	1.0	5463.000	5.283
563	1.0	7290.250	6.268
573	1.0	6388.625	5.805
583	1.0	6685.125	5.503
593	1.0	7860.875	4.517

2.2. Virtual sample generation

2.2.1. Virtual sample generation using visible-domain cubic spline interpolation

Cubic spline interpolation is a piecewise interpolation method, where each segment is computed using a cubic polynomial. Given a series of known points with corresponding x and y coordinates, $[x_0, x_1, x_2, x_3 \cdots x_n]$, the interval $[x_0, x_n]$ is divided into n segments. The goal is to calculate a cubic polynomial for each segment. The cubic spline interpolation should satisfy the following conditions;

(1) Within each interval $[x_i, x_{i+1}]$, $S_i(x)$ is represented by a cubic equation.

(2) It must coincide with the known points, meaning $S_i(x) = y_i (i = 0,1,\cdots,n)$.

(3) The curve is smooth, ensuring continuity for $S(x)$, $S'(x)$, and $S''(x)$.

The general form of each cubic polynomial is written as, $y = a_i + b_i x + c_i x^2 + dx^3$, where this equation represents the cubic spline function $S_i(x)$.

2.2.2. Virtual sample generation using visible-domain polynomial extrapolation

The limitation of cubic spline interpolation is that it can only interpolate and fit data within the range of existing data. The effectiveness of this method is limited to the covered data region. It cannot generate data beyond the region of existing data. In order to extend the prediction scope of the machine learning model, the polynomial extrapolation technique is introduced to generate the data in the invisible-domain, i.e., the domain outside the scope of the training dataset defined by the input feature space and output values. In general, machine learning predictions for unknown data sets are limited to

the range of existing training data sets. To overcome this limitation, we use techniques such as polynomial extrapolation to generate sample points that are beyond the scope of the training data. These additional sample points help to evaluate the performance of the model in a larger range, improving the ability to generalize and adapt of the model. In the polynomial extrapolation, by using the existing data for polynomial fitting, and then extrapolating based on the fitted polynomial function, the sample points beyond the range of training data are generated. Such a data enhancement approach helps the model to understand the data features more comprehensively, but attention needs to be paid to the accuracy and applicability of the extrapolated data.

2.3. NSGA-II algorithm

In 2002, Deb et al. proposed the NSGA-II algorithm [32]. The specific working flow chart of the NSGA-II algorithm is shown in Figure 2. Iterate to compute the two parameter values for each individual, the individual of $n_i = 0$ is assigned a front 1, and it is stored in the $\mathcal{F}_1$ set. Empty the middle collection Q after you have finished storing the previous collection. The $n_i - 1$ operation is performed on the S set of each individual in the $\mathcal{F}_1$ set, and the individual satisfying $n_i = 0$ is assigned to the front 2, and it is stored in the middle set Q, and the Q set is stored in the whole $\mathcal{F}_2$ set after the storage is completed. The details are shown in Algorithm 1.

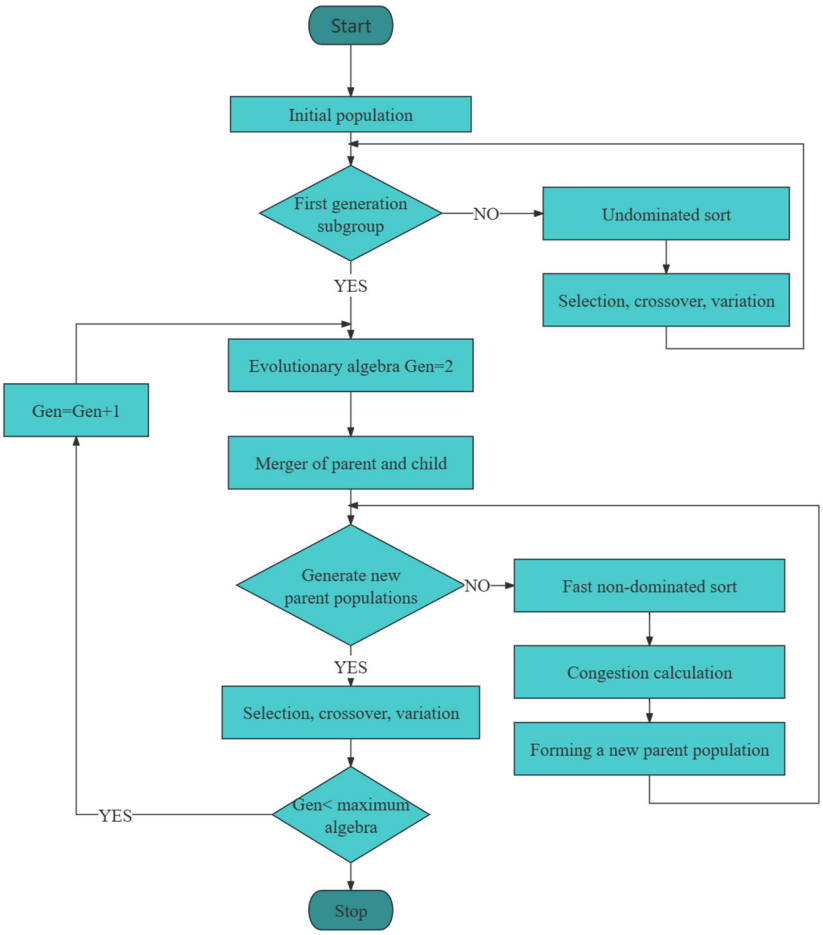

Figure 2. Flowchart of NSGA-II algorithm.

Algorithm 1. Fast-non-dominated-sort (P).

for each $p \in P$
  $S_p = \varnothing$
  $n_p = 0$
  for each $q \in P$
    if ($p \prec q$) then // If p dominates q
      $S_p = S_p \cup \{q\}$ // Add q to the set of solutions dominated by p
    else if ($q \prec p$) the
      $n_p = n_p + 1$ // Increment the domination counter of p
  if $n_p = 0$ then // p belongs to the first front
    $p_{rank} = 1$
    $\mathcal{F}_1 = \mathcal{F}_1 \cup \{p\}$
$i = 1$ // Initialize the front counter
while $\mathcal{F}_i \neq 0$
  $Q = \varnothing$ // Used to store the members of the next front
  for each $p \in \mathcal{F}_i$
    for each $q \in S_p$
      $n_q = n_q - 1$
      if $n_q = 0$ then // q belongs to the next front
        $q_{rank} = i + 1$
        $Q = Q \cup \{q\}$
  $i = i + 1$
  $\mathcal{F}_i$=Q

As shown in Algorithm 2, in each objective function, the non-dominated set l is sorted according to the function value, after which the distance of the boundary points is set to infinity, and the crowding distance of the *i-th* solution is equal to the sum of the normalized function value interpolations of the adjacent points in all objective functions.

Algorithm 2. Crowding-distance-assignment(T).

$l = |\mathcal{T}|$ // number of solutions in $\mathcal{T}$
for each i, set $\mathcal{T}[i]_{distance} = 0$ // initialize distance
for each objective m
$\mathcal{T} = sort(\mathcal{T}, m)$ // sort using each objective value
$\mathcal{T}[1]_{distance} = \mathcal{T}[l]_{distance} = \infty$ // so that boundary points are always selected
  for $i = 2$ *to* $(l-1)$ // for all other points
    $\mathcal{T}[i]_{distance} = \mathcal{T}[i]_{distance} + (\mathcal{T}[i+1] \cdot m - \mathcal{T}[i-1] \cdot m) / (f_m^{\max} - f_m^{\min})$

3. Results and discussion

3.1. Data augmentation

The small dataset size is a critical bottleneck for using machine learning methods to predict material properties. To solve the small dataset problem, data augmentation is performed based on the original dataset [33–35]. We introduce a virtual sample generation algorithm using visible-domain cubic spline interpolation and invisible-domain polynomial extrapolation to overcome the lack of training samples in this study.

Firstly, cubic spline is used for interpolation in the visible domain, generating 166 virtual samples in the visible domain. Next, we apply polynomial fitting to the original data for coercivity (H_{cj}) and maximum magnetic energy product ((BH)$_{max}$). In addition, polynomial extrapolation is used to generate 15 virtual samples in the invisible domain. In total, 181 data points are generated for the heat treatment parameters-performance of $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$.

During the data augmentation in the invisible domain**,** the original 16 data points are subjected to Min-Max normalization using the following formula:

$$X_{ij}^{'} = \frac{X_{ij} - X_{\min}^{i}}{X_{\max}^{i} - X_{\min}^{i}}, \tag{1}$$

where X_{ij} represents the value of the *j-th* variable of the *i-th* attribute, $X_{\min}^{i}$ and $X_{\max}^{i}$ correspond to the minimum and maximum values of the *j-th* attribute, respectively. $X_{ij}^{'}$ represents the normalized value of the *j-th* variable of the *i-th* attribute.

Next, we perform polynomial fitting separately on the normalized data for coercivity (H_{cj}) and maximum magnetic energy product ((BH)$_{max}$). For different ML models, the decision coefficient (R^2) and root mean square error (*RMSE*) are selected as evaluation indexes. The larger R^2 and the smaller *RMSE* mean the higher prediction accuracy of the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\overline{y} - y_i)^2}, \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}, \tag{3}$$

where n is the sample size, $\hat{y}_i$ is the predicted value, y is the real value, and $\overline{y}$ is the mean value.

The results are shown in Table 2, where the polynomial fitting for H_{cj} shows an R^2 of 0.916 for the fourth-degree polynomial. The polynomial fitting for (BH)$_{max}$ shows an R^2 of 0.937 for the third-degree polynomial and 0.949 for the fourth-degree polynomial. Considering the impact of model complexity, we ultimately select fourth-degree polynomial regression for coercivity (H_{cj}) and a third-degree polynomial model for (BH)$_{max}$. The specific expressions are as follows:

$$\begin{aligned} y_1 = &-0.885 + 7.098\times x_1 + 1.533\times x_2 - 18.352\times x_1^2 - 9.027\times x_1\times x_2 + 3.314\times x_2^2 \\ &+37.981\times x_1^3 - 5.936\times x_1^2\times x_2 - 1.768\times x_1\times x_2^2 + 0.323\times x_2^3 - 27.032\times x_1^4 \\ &+20.981\times x_1^3\times x_2 - 16.163\times x_1^2\times x_2^2 + 12.945\times x_1\times x_2^3 - 4.277\times x_2^4, \end{aligned} \tag{4}$$

$$y_2 = 1.137 - 4.770 \times x_1 - 2.080 \times x_2 + 2.837 \times x_1^2 + 11.334 \times x_1 \times x_2 \\ -1.731 \times x_2^2 - 0.244 \times x_1^3 - 4.101 \times x_1^2 \times x_2 - 5.671 \times x_1 \times x_2^2 + 3.395 \times x_2^3, \quad (5)$$

where x_1 is the annealing temperature, x_2 is the magnetic field strength, y_1 is the coercive force, and y_2 is the maximum magnetic energy product.

Table 2. The results of quadratic, cubic, and quartic polynomial fittings for coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) are presented.

Polynomial	$H_{cj}(Oe)$		$(BH)_{max}(MGOe)$	
	R^2	*RMSE*	R^2	*RMSE*
Quadratic Polynomial	0.219	0.234	0.780	0.133
Cubic Polynomial	0.466	0.193	0.937	0.075
Quartic Polynomial	0.916	0.077	0.949	0.067

Using the established polynomial models, we conduct polynomial extrapolation for H_{cj} and $(BH)_{max}$, generating 15 virtual samples in the invisible domain. The fitting results of the polynomial models for H_{cj} and $(BH)_{max}$ are shown in Figure 3.

3.2. Data validation

Before training the model using the generated data, it is necessary to validate whether the virtual samples generated are reasonable. This study mainly explains the validation process based on model performance and visualization. Firstly, the generated data is mixed with the original data, and an ML model is trained using ten-fold cross-validation. Next, scatter plots and violin plots are used to visualize the original dataset and the virtual sample dataset. Finally, the R^2 value of the model is compared before and after data augmentation to further prove the validity and usability of the generated virtual samples.

To validate the reasonableness of the generated data, Figure 4(a),(b) depict the coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) for both virtual and real samples, respectively. These plots show that the distribution of generated virtual samples is similar to that of the original samples. Figure 5(a) and Figure 5(b) display the scatter plots of the spatial distribution of H_{cj} and $(BH)_{max}$, respectively. Figure 5(c) presents the scatter plot depicting the distribution of generated virtual samples alongside real samples, where the inclusion of virtual samples leads to a more uniform distribution and reduced differences between data values. Through ten-fold cross-validation on the generated virtual samples, the results presented in Figure 6 indicate that the enhanced dataset, which incorporates cubic spline interpolation and polynomial extrapolation, performs better in training models compared to the original 16 data points, where we refer to ML models XGBoost, RF, and LightGBM after data augmentation as DA_XGBoost, DA_RF, and DA_LightGBM, respectively. Among these models, DA_XGBoost stands out as the most effective, yielding R^2 values of 0.922 for coercivity (H_{cj}) and 0.983 for maximum magnetic energy product ($(BH)_{max}$). These R^2 values are significantly higher than those obtained from the original dataset, demonstrating a substantial enhancement in predicting coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$). This observation underscores the efficacy of the methodology proposed in this study for data enhancement.

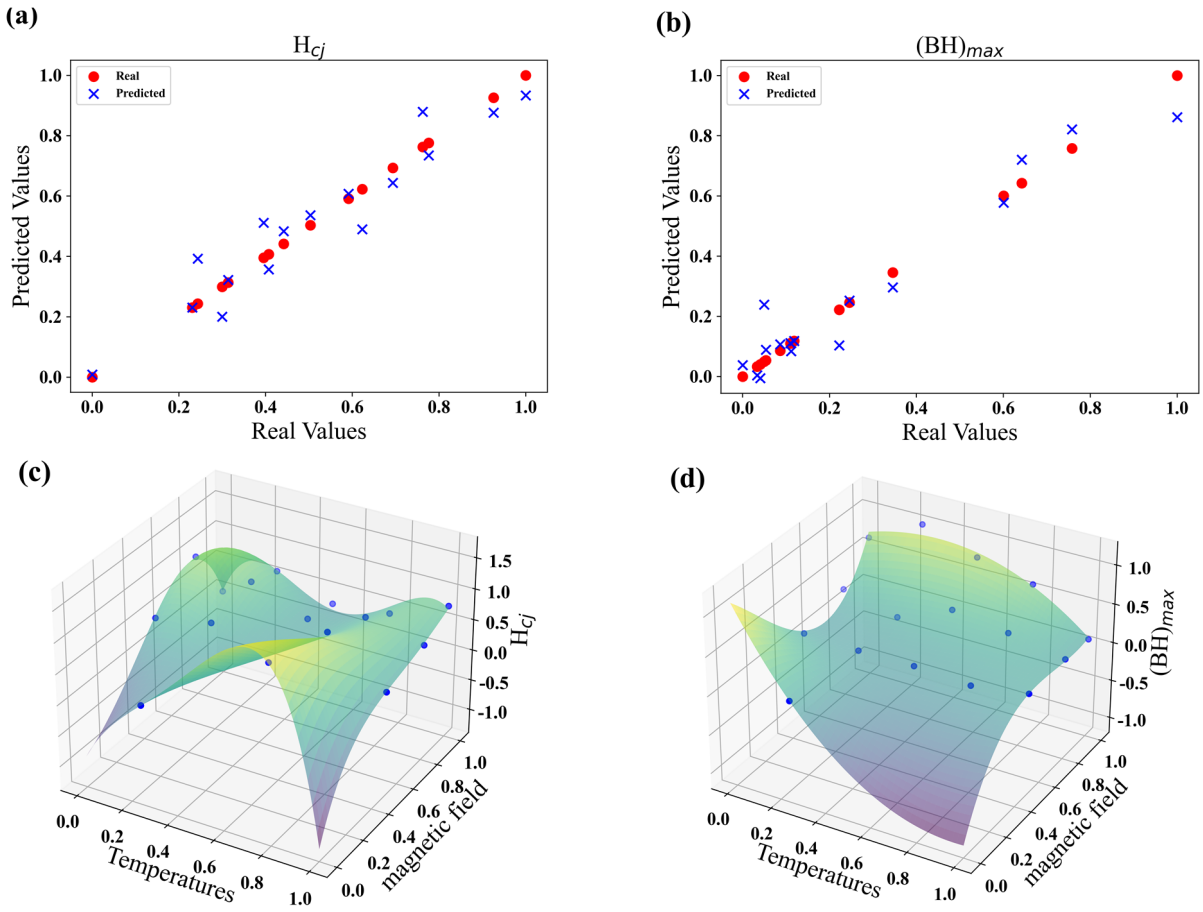

Figure 3. The results of polynomial regression for coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) in relation to the heat treatment parameters of 16 samples of $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$. (a) Fourth-degree polynomial regression for coercivity (H_{cj}). (b) Third-degree polynomial regression for maximum magnetic energy product ($(BH)_{max}$). (c) Surface plot of fourth-degree polynomial fitting for coercivity (H_{cj}). (d) Surface plot of third-degree polynomial fitting for maximum magnetic energy product ($(BH)_{max}$).

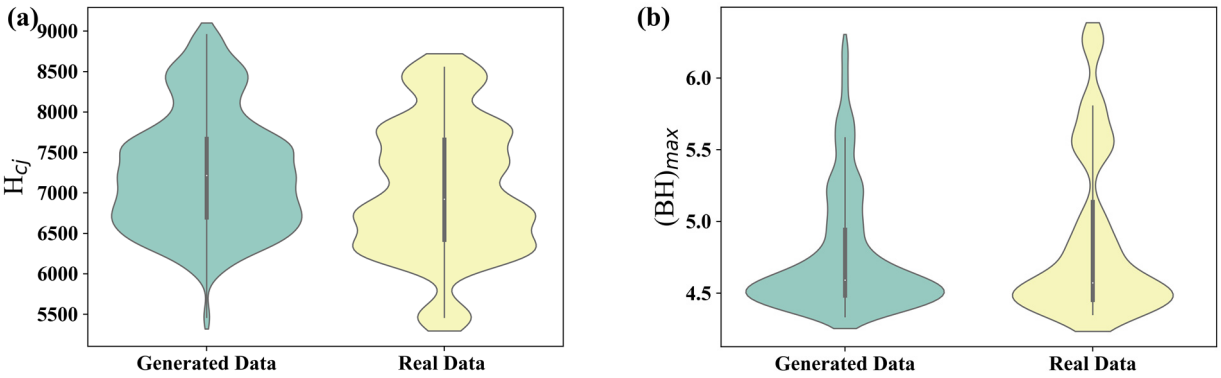

Figure 4. (a) Violin plots comparing the generated virtual data and real data for coercivity (H_{cj}). (b) Violin plots comparing the generated virtual data and real data for maximum magnetic energy product ($(BH)_{max}$).

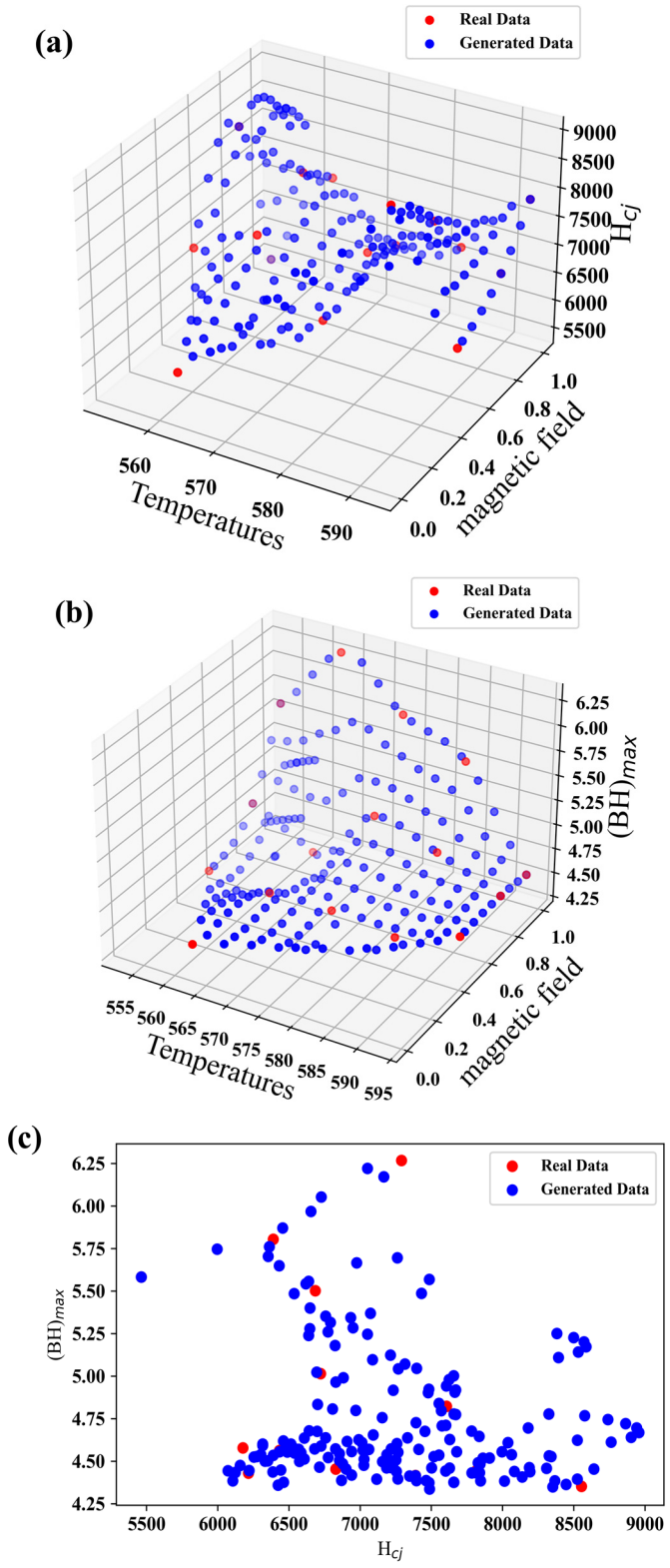

Figure 5. (a) Scatter plot distribution in parameter space for coercivity (H_{cj}) comparing generated virtual data and real data. (b) Scatter plot distribution in parameter space for maximum magnetic energy product ($(BH)_{max}$) comparing generated virtual data and real data. (c) Two-dimensional scatter plot comparing the generated virtual data and real data for coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$).

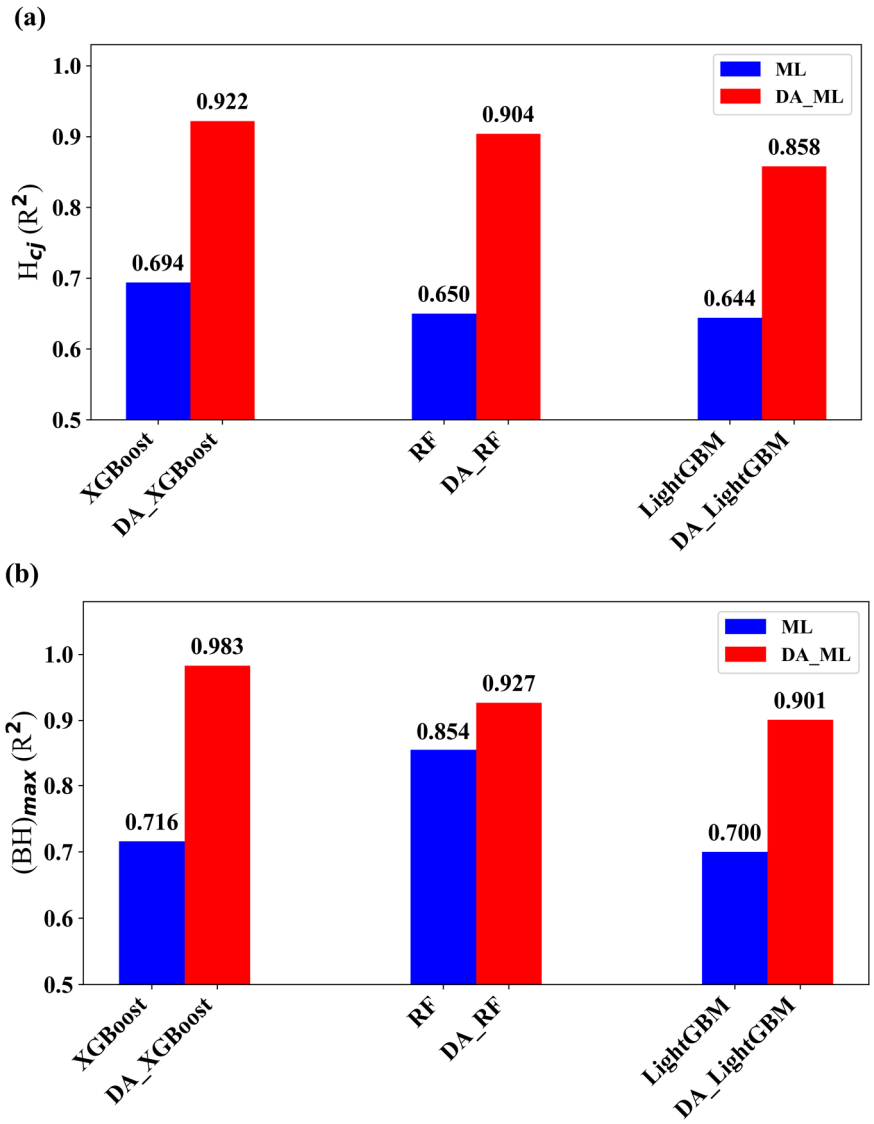

Figure 6. (a)The R^2 of ML models for the prediction of H_{cj} before and after data augmentation. (b)The R^2 of models for the prediction of $(BH)_{max}$ before and after data augmentation, where DA_XGBoost, DA_RF, and DA_LightGBM are the models trained after data augmentation.

3.3. Model construction

The purpose of machine learning modeling is to establish the relationship between the magnetic properties of materials and heat treatment process parameters. Considering the benefits of ensemble learning, which combines multiple learners to improve performance and generalization, this study establishes Random Forest, XGBoost, and LightGBM models to predict coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) of rare earth hard magnetic materials, respectively.

The models are trained using the generated virtual samples and tested on the real sample dataset. The results are shown in Table 3. Figure 7 is a comparison of the models for R^2. Both DA_XGBoost and DA_RF models achieved an R^2 of above 0.95 for coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$). The slightly lower performance of the DA_LightGBM model may be due to the

relatively small size of the dataset, as this study generated a total of 181 virtual samples. DA_LightGBM typically excels in large-scale datasets but may exhibit average performance on small-scale datasets. Among them, DA_XGBoost performed the best, and the performance of the model on the training dataset and real dataset is shown in Figure 8. Therefore, for subsequent research, the selected model will be DA_XGBoost.

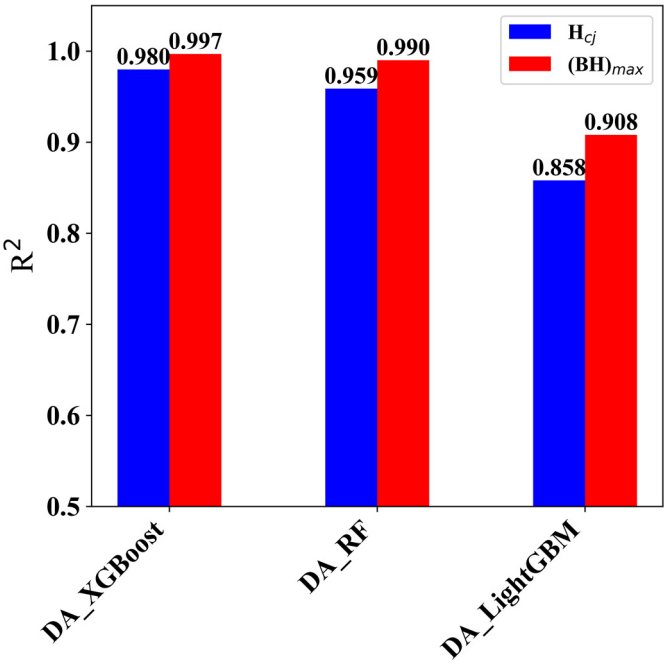

Figure 7. The R^2 of DA_XGBoost, DA_RF, and DA_LightGBM models on tested real datasets for H_{cj} and $(BH)_{max}$.

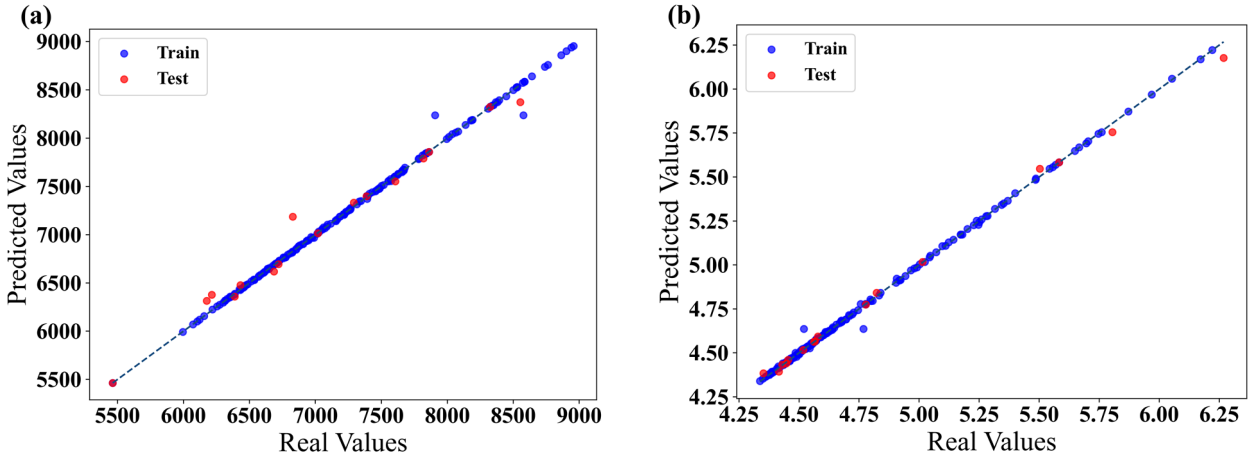

Figure 8. The predictive performance of DA_XGBoost on the training dataset and real dataset of 16 samples is shown. (a) Coercivity prediction (H_{cj}) of DA_XGBoost model after data enhancement. (b) Prediction of maximum magnetic energy product ($(BH)_{max}$) by data-enhanced DA_XGBoost model.

Table 3. The performance of different prediction models for H_{cj} and $(BH)_{max}$ is evaluated by training on the generated dataset and testing on the real dataset.

Model	$H_{cj}(Oe)$		$(BH)_{max}(MGOe)$	
	R^2	*RMSE*	R^2	*RMSE*
DA_XGBoost	0.980	116.996	0.997	0.031
DA_RF	0.959	165.496	0.990	0.057
DA_LightGBM	0.858	290.355	0.908	0.188

3.4. NSGA-II multi-Objective optimization and heat treatment parameter recommendations

For the multi-objective optimization of alloys, Xie et al. enhanced the combination of ultimate tensile strength and electric conductivity using machine learning [23]. Wei et al. designed the lead-free solder alloys with high strength and high ductility [37]. Faheem et al. optimized the parametric of a shape memory alloy using NSGA-II [38]. Based on the obtained model, we attempt to use available machine learning models and the non-dominated sorting genetic algorithm-II (NSGA-II) to find heat treatment parameters for the magnetic material $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$ that simultaneously possess higher coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$).

Using the previously trained DA_XGBoost model, we conduct NSGA-II multi-objective optimization, and the results are shown in Figure 9. The model recommended 12 sets of heat treatment parameters, including annealing temperature and magnetic field strength, as well as their corresponding coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$). The specific values of coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) calculated using the previously trained polynomial regression (Poly) are provided in Table 4. From Table 6, we can observe that the values of coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) obtained through NSGA-II multi-objective optimization are very close to those calculated using the polynomial models. This further demonstrates the reliability of the heat treatment parameters recommended in this study.

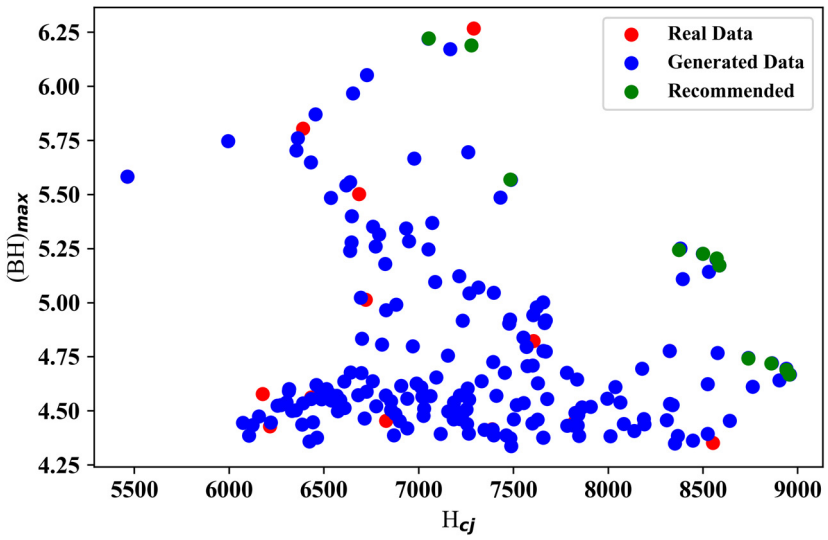

Figure 9. The optimal combined distribution of coercivity (H_{cj}) and maximum magnetic energy product of $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$ is represented by green dots.

Table 4. Presents the top 12 heat treatment parameter combinations of $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$, ranked by high coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$), along with corresponding predictions of coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) using both DA_XGBoost and polynomial regression (Poly).

Material	Temperatures （K）	magnetic field （T）	DA_XGBoost H_{cj}	DA_XGBoost $(BH)_{max}$	Poly H_{cj}	Poly $(BH)_{max}$
1	557.00	0.80	8955.11	4.67	8957.90	4.67
2	563.67	0.93	7483.63	5.57	7769.78	5.51
3	560.99	0.90	8373.09	5.24	8383.57	5.25
4	563.67	1.00	7277.33	6.19	7252.93	6.01
5	566.33	1.00	7052.58	6.22	6842.33	6.02
6	559.02	0.80	8860.06	4.72	8861.67	4.73
7	557.99	0.90	8585.94	5.17	8585.98	5.17
8	559.00	0.90	8572.83	5.20	8571.06	5.20
9	558.00	0.80	8940.31	4.69	8940.35	4.70
10	560.00	0.90	8499.49	5.23	8499.18	5.23
11	566.33	1.00	7052.58	6.22	6842.33	6.02
12	560.00	0.80	8739.10	4.74	8738.72	4.75

4. Conclusions

In conclusion, we proposed a new framework for multi-objective optimization of processing parameters using machine learning models to predict coercivity (H_{cj}) and maximum magnetic energy product ($(BH)_{max}$) for rare-earth hard magnetic materials. Moving on to predicting the properties of $Nd_7Fe_{67.5}B_{21}Nb_{2.5}Zr_2$ based on heat treatment parameters, it encountered the challenge of limited experimental data, with only 16 relevant samples available. To address this, we introduced virtual sample generation methods based on visible-domain cubic spline interpolation and invisible-domain polynomial extrapolation, expanding the dataset to 197 samples. Ensemble learning is used to train the dataset after data enhancement. Among them, DA_XGBoost model has the highest prediction accuracy for H_{cj} and $(BH)_{max}$ on the real data set, with R^2 0.980 and 0.997, respectively. Through the multi-objective optimization of NSGA-II in the virtual heat treatment parameter space, 12 sets of processing parameters for magnetic materials with high coercivity and maximum magnetic energy product are recommended, and their heat treatment parameters are better than those of existing rare earth hard magnetic materials. Among them, the coercive force and maximum magnetic energy product of the optimized magnet can be increased by 4.7% and 7.4%, respectively.

In summary, the present work provides efficient and accurate ML models for predicting H_{cj} and $(BH)_{max}$ of rare-earth hard magnetic materials based on their processing parameters. This provides new insights and methods for the design of rare earth hard magnetic materials, with the potential to drive performance improvements in practical applications to meet various performance requirements. Additionally, we offer a design framework for simultaneously enhancing H_{cj} and $(BH)_{max}$ in rare-earth hard magnetic materials, contributing to time and cost savings in the exploration of high-performance magnetic materials and accelerating the development and design of rare-earth magnetic materials. This study also provides a new framework for the design of other materials with multi-objective performances.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by the National Key R&D Program of China (2024YFB3813701).

Conflict of interest

The authors declare that they have no conflicts of interest with respect to the research, authorship, and publication of the article.

References

1. Gutfleisch O, Willard MA, Brück E, Chen CH, Sankar SG, Liu JP, (2011) Magnetic materials and devices for the 21st century: Stronger, lighter, and more energy efficient. *Adv Mater* 23: 821–42. https://doi.org/10.1002/adma.201002180
2. Chen J, J Liu, MJ Zhang, ZJ Dong, ZJ Peng, XY Ji, et al. (2023) Accelerated discovery of cost-effective Nd-Fe-B magnets through adaptive learning. *J Mater Chem A* 16: 8988–9001. https://doi.org/10.1039/d2ta10043f
3. Hono K, Sepehri-Amin H, (2018) Prospect for HRE-free high coercivity Nd-Fe-B permanent magnets. *Scr Mater* 151: 6–13. https://doi.org/10.1016/j.scriptamat.2018.03.012
4. AK Pathak, M Khan, K Gschneidner, R McCallum, L Zhou, K Sun, et al. (2016) Magnetic properties of bulk, and rapidly solidified nanostructured $(Nd_{1-x}Ce_x)_2Fe_{14-y}Co_yB$ ribbons. *Acta Mater* 103: 211–216. https://doi.org/10.1016/J.ACTAMAT.2015.09.049
5. Huang QF, Jiang QZ, Shi Y, Rehman SU, Wei X, Li ZX, et al. (2022) Enormous improvement of the coercivity of Ga and Cu co-doping Nd-Fe-B sintered magnet by post-sinter annealing. *J Alloys Compd* 894: 162418. https://doi.org/10.1016/j.jallcom.2021.162418
6. Baldissera AB, Pavez P, Wendhausen PAP, Ahrens CH, Mascheroni JM, (2017) Additive manufacturing of bonded Nd-Fe-B—Effect of process parameters on magnetic properties. *IEEE Trans Magn* 53: 1–4. https://doi.org/10.1109/TMAG.2017.2715722
7. Vial F, Joly F, Nevalainen E, Sagawa M, Hiraga K, Park KT, (2002) Improvement of coercivity of sintered NdFeB permanent magnets by heat treatment. *J Magn Magn Mater* 242: 1329–1334. https://doi.org/10.1016/S0304-8853%2801%2900967-2
8. Zhang TQ, Liu SS, Wu ZW, Zhang LT, Yu R, (2022) Coercivity and remanence enhancement in hot-deformed Nd-Fe-B magnets by high-temperature short-term annealing process. *J Alloys Compd* 903: 163975. https://doi.org/10.1016/j.jallcom.2022.163975
9. Tan XH, Tan YF, Tang YJ, Sun SB, Zhang SY, Xu H, (2021) The improvement of magnetic property by grain refinement using magnetic field annealing crystalline $(Nd_{0.8}Pr_{0.2})_{2.2}Fe_{12}Co_2B$ ribbons. *J Magn Magn Mater* 518: 167434. https://doi.org/10.1016/j.jmmm.2020.167434
10. Jia P, Liu JM, Wang EG, Han K, (2013) The effects of high magnetic field on crystallization of $Fe_{71}(Nb_{0.8}Zr_{0.2})_6B_{23}$ bulk metallic glass. *J Alloys Compd* 581: 373–377. https://doi.org/10.1016/J.JALLCOM.2013.07.066

11. Gu Y, Li Z, Tan XH, Pan MX, Xu H, (2018) Effect of magnetic field treatment on magnetic properties and microstructure of $Fe_{69.5}Nd_7B_{21}Nb_{2.5}$ bulk permanent magnets. *Phys B* 546: 15–20. https://doi.org/10.1016/J.PHYSB.2018.07.011
12. Qiao ZQ, Dong SZ, Li Q, Lu XM, Chen RJ, Guo S, et al. (2023) Performance prediction models for NdFeB using machine learning methods and interpretable studies. *J Alloys Compd* 963: 171250. https://doi.org/10.1016/j.jallcom.2023.171250
13. Herbst JF, (1991) $R_2Fe_{14}B$ materials: Intrinsic properties and technological aspects. *Rev Mod Phys* 63: 819. https://doi.org/10.1103/RevModPhys.63.819
14. Delette G, (2023) $Nd_2Fe_{14}B$ permanent magnets substituted with non-critical light rare earth elements (Ce, La): A review. *J Magn Magn Mater* 577: 170768. https://doi.org/10.1016/j.jmmm.2023.170768
15. Yu LP, Guo XX, Wang G, Sun BA, Han DX, Chen C, et al. (2022) Extracting governing system for the plastic deformation of metallic glasses using machine learning. *Sci China Phys Mech Astron* 65: 264611. https://doi.org/10.1007/s11433-021-1840-9
16. Jie X, Zhang TY, Shi SQ, (2020) Machine learning of mechanical properties of ateels. *Sci China Technol Sci* 63: 1247–1255. https://doi.org/10.1007/s11431-020-1599-5
17. Bertolucci CL, Torres D, Vangrunderbeek V, Bernal M, Paldino GM, Bontempi G, et al. (2023) Estimating pitting descriptors of 316 L stainless steel by machine learning and statistical analysis. *NPJ Mate Degrad* 7: 82. https://doi.org/10.1038/s41529-023-00403-z
18. Shen C, Wang C, Rivera-Díaz-del-Castillo PEJ, Xu D, Zhang Q, Zhang C, et al. (2021) Discovery of marageing steels: Machine learning vs. physical metallurgical modelling. *J Mater Sci Technol* 87: 258–268. https://doi.org/10.1016/J.JMST.2021.02.017
19. Zhang M, Sun CN, Zhang X, Goh PC, Wei J, Hardacre D, et al. (2019) High cycle fatigue life prediction of laser additive manufactured stainless steel: A machine learning approach. *Int J Fatigue* 128: 105194. https://doi.org/10.1016/J.IJFATIGUE.2019.105194
20. Lee JY, Kim M, Lee YK, (2022) Design of high strength medium-Mn steel using machine learning. *Mater Sci Eng A* 843: 143148. https://doi.org/10.1016/j.msea.2022.143148
21. Chen C, Zang DD, Ren JL, (2025) A new multi-objective optimization design framework for high-temperature alloy steel via machine learning. *Materials Today Communications* 49: 2352-4928. https://doi.org/10.1016/j.mtcomm.2025.113955.
22. Rao ZY, Tung PY, Xie RW, Wei Y, Zhang H, Ferrari A, et al, (2022) Machine learning-enabled high-entropy alloy discovery. *Science* 378 78–85. https://doi.org/10.1126/science.abo4940
23. Zhang HT, Fu HD, He XQ, Wang C, Jiang L, Chen LQ, et al. (2020) Dramatically enhanced combination of ultimate tensile strength and electric conductivity of alloys via machine learning screening. *Acta Mater* 200: 803–810. https://doi.org/10.2139/ssrn.3646448
24. Chen C, Ma LY, Zhang Y, Liaw PK, Ren JL, (2023) Accelerating the design of high-entropy alloys with high hardness by machine learning based on particle swarm optimization. *Intermetallics* 154: 107819. https://doi.org/10.1016/j.intermet.2022.107819
25. Chen C, Zhou HR, Long WM, Wang G, Ren JL, (2023) Phase prediction for high-entropy alloys using generative adversarial network and active learning based on small datasets. *Sci China Tech Sci* 66: 3615–3627. https://doi.org/10.1007/s11431-023-2399-2
26. Wei QH, Cao B, Deng L, Sun A, Dong Z, Zhang TY, (2023) Discovering a formula for the high temperature oxidation behavior of FeCrAlCoNi based high entropy alloys by domain knowledge-guided machine learning. *J Mater Sci Technol* 149: 237–246. https://doi.org/10.1016/j.jmst.2022.11.040

27. Chen C, Han XL, Zhang Y, Peter K. Liaw, Ren JL, (2024) Phase prediction of high-entropy alloys based on machine learning and an improved information fusion approach. *Computational Materials Science* 239: 112976. https://doi.org/10.1016/j.commatsci.2024.112976.
28. Gu Y, Li Z, Wang CX, Wu Z, Feng X, Zhang Y, et al. (2020) Enhancing hard magnetic properties in $Fe_{69.5-x}Nd_7B_{21}Nb_{2.5}Zr_x$ (x=0–3) bulk magnets by magnetic field heat treatment. *J Magn Magn Mater* 499: 166215. https://doi.org/10.1016/j.jmmm.2019.166215
29. Tan X, Tan Y, Tang Y, Sun S, Zhang S, Xu H, (2021) The improvement of magnetic property by grain refinement using magnetic field annealing crystalline $(Nd_{0.8}Pr_{0.2})_{2.2}Fe_{12}Co_2B$ ribbons. *J Magn Magn Mater* 518: 167434. https://doi.org/10.1016/j.jmmm.2020.167434
30. Wang YH, Tian YF, Kirk T, Laris O, Ross JH, Noebe RD, et al. (2020) Accelerated design of Fe-based soft magnetic materials using machine learning and stochastic optimization. *Acta Mater* 194: 144–155. https://doi.org/10.1016/j.actamat.2020.05.006
31. Lambard G, Sasaki TT, Sodeyama K, Ohkubo T, Hono K, (2022) Optimization of direct extrusion process for Nd-Fe-B magnets using active learning assisted by machine learning and Bayesian optimization. *Scripta Mater* 209: 114341. https://doi.org/10.1016/j.scriptamat.2021.114341
32. Deb K, Pratap A, Agarwal S, Meyarivan T, (2002) A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* 6: 182–197. https://doi.org/10.1109/4235.996017
33. Kim Y, Kim Y, Yang C, Park K, Gu GX, Ryu S, (2021) Deep learning framework for material design space exploration using active transfer learning and data augmentation. *NPJ Comput Mater* 7: 140. https://doi.org/10.1038/s41524-021-00609-2
34. Sutojo T, Rustad S, Akrom M, Syukur A, Shidik GF, Dipojono HK, (2023) A machine learning approach for corrosion small datasets. *NPJ Mater Degrad* 7: 18–10. https://doi.org/10.1038/s41529-023-00336-7
35. Xu PC, Ji XB, Li MJ, Lu WC, (2023) Small data machine learning in materials science. *NPJ Comput Mater* 9: 42. https://doi.org/10.1038/s41524-023-01000-z
36. Hua Y, Zhu H, Xu Y, (2020) Multi-objective optimization design of bearingless permanent magnet synchronous generator, *IEEE Trans Appl Supercond* 30: 1–5. https://doi.org/10.1109/TASC.2020.2970661
37. Wei QF, Cao B, Yuan H, Chen Y, You K, Yu S, et al., (2023) Divide and conquer: Machine learning accelerated design of lead-free solder alloys with high strength and high ductility. *NPJ Comput Mater* 9: 201. https://doi.org/10.1038/s41524-023-01150-0
38. Faheem AB, Hasan F, Khan AA, Singh B, Ayaz M, Shamim FA, et al. (2023) Parametric optimization of electric discharge machining of $Ni_{55.65}Ti$ based shape memory alloy using NSGA II with TOPSIS. *J Mater Res Technol* 26: 1306–1324. https://doi.org/10.1016/j.jmrt.2023.07.259

AIMS Press