

Data-driven discovery of high-entropy rare earth aluminates for high-temperature thermal barrier applications

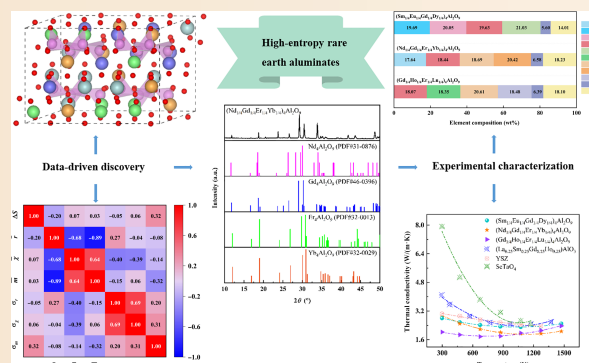
Kaili Chu¹, Wenhui Zhao¹, Yun Fan¹, Huimin Xiang^{2,✉}, Yuchen Liu³, Yiran Li^{1,4}, Wenxian Li⁵, Bin Liu^{1,4,✉}

Cite this article: Chu K, Zhao W, Fan Y, et al. *J Adv Ceram* 2026, **15**(9): 9221354. <https://doi.org/10.26599/JAC.2026.9221354>

ABSTRACT: High-entropy rare earth (RE) aluminates are promising candidates for thermal/environmental barrier coatings (T/EBCs), while the immense compositional space presents significant challenges for traditional experimental discovery. To address this issue, artificial neural network (ANN), support vector machine (SVM), and random forest classification (RFC) are employed as three machine learning models to predict the single-phase formation ability of $(\text{RE}^{I}_{1/4}\text{RE}^{II}_{1/4}\text{RE}^{III}_{1/4}\text{RE}^{IV}_{1/4})\text{Al}_2\text{O}_9$ materials. Both the ANN and RFC models achieve optimal validation accuracy, demonstrating their outstanding ability to capture complex patterns from the dataset. SHapley Additive exPlanations (SHAP) analysis is utilized to interpret the contribution of feature descriptors, revealing the significant

impact of factors such as average ionic radius ($\bar{r}$) on phase stability. According to the prediction results of machine learning, three representative ceramic samples are selected, and single-phase monoclinic crystal structures and uniform elemental distribution are confirmed by X-ray diffraction (XRD) and scanning electron microscopy (SEM). The synthesized ceramics exhibit quasiductile behavior with enhanced damage tolerance combined with lower thermal conductivity, thus making them promising candidates for next-generation T/EBCs.

KEYWORDS: high-entropy rare earth (RE) aluminates; machine learning; single-phase formation ability; thermal/environmental barrier coatings (T/EBCs)



1 Introduction

High-entropy ceramics were first demonstrated by Rost *et al.* [1] in 2015 with the synthesis of a rock-salt-structured (MgNiCuCoZn)O ceramic, expanding the high-entropy concept from the field of alloys to ceramic systems. High-entropy ceramics represent a class of single-phase solid solutions in which four or more elements occupy equivalent sites within the crystal lattice. The vast composition design space and complex atomic structure of high-entropy ceramics endow them with the high-entropy effect, lattice distortion effect, sluggish diffusion effect, and cocktail effect [2,3]. In comparison with the behavior of single-component compounds, high-entropy ceramics have superior high-temperature stability, slower grain growth rate, better mechanical and thermal properties, as well as better environmental resistance [4]. These outstanding properties render high-entropy ceramics well positioned to meet the demands of diverse applications, including thermal/environmental barrier coatings (T/EBCs) [5], electromagnetic interference shielding [6], high-end manufacturing [7], thermoelectrics [8], and catalysts [9].

In the context of T/EBCs, the demand for materials with high-temperature phase stability, low thermal conductivity, and excellent sintering resistance at elevated temperatures is becoming increasingly stringent, driving the exploration of novel ceramic systems beyond conventional yttria-stabilized zirconia (YSZ) [10,11]. High-entropy ceramics offer a promising pathway to address these challenges, as the severe lattice distortion caused by the occupation of multiprincipal elements effectively enhances phonon scattering, thereby reducing thermal conductivity [12]. Meanwhile, the high-entropy effect expands the stability region of single-phase solid solutions, which not only inhibits phase transformations and hinders grain coarsening under extreme thermal gradients but also provides superior CaO–MgO–Al₂O₃–SiO₂ (CMAS) resistance. For instance, Liu *et al.* [13] found that $(\text{La}_{0.25}\text{Sm}_{0.25}\text{Gd}_{0.25}\text{Ho}_{0.25})\text{AlO}_3$ shows a reduced thermal conductivity of 4.11 W/(m·K) at room temperature and thermal expansion coefficients of $10.10 \times 10^{-6} \text{ K}^{-1}$ (25–1300 °C). After annealing at 1590 °C for 18 h, the average grain size of $(\text{Y}_{0.2}\text{Yb}_{0.2}\text{Lu}_{0.2}\text{Eu}_{0.2}\text{Er}_{0.2})_3\text{Al}_5\text{O}_{12}$ increases only from 1.56 to 2.27 μm

¹School of Materials Science and Engineering, Shanghai University, Shanghai 200444, China. ²School of Materials Science and Engineering, Zhengzhou University, Zhengzhou 450001, China. ³College of Sciences, Nanjing Agricultural University, Nanjing 210095, China. ⁴State Key Laboratory of Advanced Refractories, Shanghai University, Shanghai 200444, China. ⁵School of Chemical Engineering, The University of New South Wales, Sydney 2052, Australia.

✉ Corresponding authors. E-mail: H. Xiang, hmxiang@alum.imr.ac.cn; B. Liu, binliu@shu.edu.cn

Received: May 28, 2026; Revised: July 24, 2026; Accepted: July 26, 2026

© The Author(s) 2026. This is an open access article under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0, <http://creativecommons.org/licenses/by/4.0/>).

[14]. Moreover, $(\text{Nd}_{0.2}\text{Sm}_{0.2}\text{Eu}_{0.2}\text{Y}_{0.2}\text{Yb}_{0.2})_4\text{Al}_2\text{O}_9$ exhibits good high-temperature phase stability, and no phase transition is detected from 300 to 1473 K [15]. Compared to their single-component counterparts, high-entropy $(\text{Sc}_{0.2}\text{Lu}_{0.2}\text{Yb}_{0.2}\text{Er}_{0.2}\text{Y}_{0.2})\text{PO}_4$ forms the thinnest reaction layer after 96 h of CMAS corrosion because of the slower mass transport of the mixed rare earth (RE) elements [16]. These intrinsic advantages make high-entropy ceramics particularly attractive for next-generation T/EBCs. Notably, high-entropy RE aluminates have emerged as an attractive class of T/EBCs due to their high melting points, good damage tolerance, and tunable thermal expansion coefficients. Despite the rapid progress realized in high-entropy aluminates, the development of new materials still relies heavily on traditional trial-and-error methods, which readily lead to time-consuming and labor-intensive tasks. Furthermore, the scarcity of research on high-entropy $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})_4\text{Al}_2\text{O}_9$ to date has considerably limited their development.

To efficiently screen the optimal combinations of single-phase high-entropy ceramics that meet specific requirements, high-throughput density functional theory (DFT) [17], molecular dynamics (MD) [18], and phase diagram calculations (CALPHAD) [19] have been applied to predict the phases of high-entropy ceramics. Computational simulations hold a distinct advantage over experiments in providing rapid and high-precision predictions. However, their weakness lies in the fact that the computational complexity grows exponentially with the scale of the system. This prohibitive cost severely restricts their applicability in exploring and predicting materials across broad composition ranges. In recent years, data-driven machine learning has been widely recognized as a groundbreaking tool in materials engineering, accelerating the discovery, design, and performance prediction of novel materials [20,21]. Drawing from the provided dataset, machine learning can uncover the underlying connections between critical parameters and target attributes [22]. For example, Pilania *et al.* [23] proposed a systematic feature engineering method and a robust learning framework that can rapidly and accurately predict the electronic bandgaps of double perovskites. The successful application of machine learning in complex material spaces has also facilitated its use in predicting the phase stability of high-entropy ceramics. Zhang *et al.* [24] used well-trained machine learning models to estimate the single-phase probability of 90 high-entropy carbide ceramics that have not yet been reported in experimental studies. Additionally, Fan *et al.* [25] successfully predicted the single-phase formation ability of 84 $(\text{RE}^{\text{I}}_{1/6}\text{RE}^{\text{II}}_{1/6}\text{RE}^{\text{III}}_{1/6}\text{RE}^{\text{IV}}_{1/6}\text{Gd}_{1/6}\text{Yb}_{1/6}\text{Lu}_{1/6})_2\text{Si}_2\text{O}_7$ compounds employing two decision fusion methods that synergize Bayesian theory and majority voting. Despite these achievements highlighting the potential of integrated approaches to accelerate material development, the application of data-driven methods in the design of high-entropy RE aluminates remains largely unexplored.

Herein, machine learning is combined with first-principles calculations to handle the challenges posed by small datasets and data imbalance, with the aim of developing accurate and effective machine learning models to predict the phase formation ability of $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})_4\text{Al}_2\text{O}_9$. The framework primarily comprises four modules, consisting of database collection and establishment, model construction and feature selection, oriented design, and experimental validation. First, a dataset containing 54 multiprincipal RE aluminates is constructed through literature collection and experiments. Then, appropriate feature descriptors are selected to predict the phase formation ability of the material using machine learning models. The feature descriptors used in model development are low-cost features that can be queried or obtained through simple calculations to reduce model costs, such

as ionic radius difference parameters and electronegativity difference parameters. Finally, $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ are successfully synthesized, and their phase structures, microstructures, and mechanical and thermal properties are systematically characterized to evaluate their potential for high-temperature T/EBCs.

2 Methods

2.1 DFT calculation

First-principles calculations based on DFT were performed using the Vienna *Ab initio* Simulation Package (VASP) [26,27]. The projector augmented wave (PAW) [28,29] method was employed together with the Perdew–Burke–Ernzerhof for solids (PBEsol) [30] generalized gradient approximation (GGA) [31] functional. In the expansion of the electronic wave function, the kinetic energy cutoff of the plane wave basis set was chosen to be 500 eV to ensure convergence, while high precision was adopted to guarantee the accuracy of the calculation. The energy convergence threshold for the electronic self-consistent field iteration was set to 10^{-6} eV/atom. To accommodate the complex composition of multiprincipal component RE aluminates $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})_4\text{Al}_2\text{O}_9$, a supercell containing 120 atoms was constructed using the special quasirandom structure (SQS) method, as shown in Figs. 1(a) and 1(b). The advantage of the SQS method lies in its ability to fully simulate the statistical properties of random ceramics or alloys using relatively small supercells. All SQS supercells were sampled in the Brillouin zone of reciprocal space using a $4 \times 2 \times 2$ Monkhorst-Pack k -point mesh. The ion relaxation process was carried out using the conjugate gradient algorithm, with the force convergence criterion of 0.005 eV/Å. To obtain the most representative random solid solution and the optimal supercell model with more stable energy, three independent SQS supercells were generated for each multiprincipal RE aluminate. A series of volume calculations ranging from 97% to 103% was performed for each supercell, and the equilibrium volume was identified by fitting energy–volume curves based on the third-order Birch–Murnaghan equation of state (EOS) [32,33], as shown in Eq. (1):

$$E(V) = E_0 + \frac{9V_0B_0}{16} \left\{ \left[\left(\frac{V_0}{V} \right)^{\frac{2}{3}} - 1 \right]^3 B'_0 + \left[\left(\frac{V_0}{V} \right)^{\frac{2}{3}} - 1 \right]^2 \left[6 - 4 \left(\frac{V_0}{V} \right)^{\frac{2}{3}} \right] \right\} \quad (1)$$

where V_0 is the equilibrium volume, E_0 is the corresponding minimum total energy, and B_0 is the equilibrium bulk modulus. B'_0 is the derivative of the bulk modulus with respect to pressure at zero pressure. Generally, the B_0 and B'_0 values of many substances were quite stable. On the basis of our test results and experimental experience [34], it was determined that $B_0 = B'_0 = 3.5$.

To predict the formation rule of multiprincipal component solid solutions, the thermodynamic stability of the material was first considered. If the solid solution phase had the lowest mixing Gibbs free energy (ΔG_{mix}) among all the possible formed phases, it was most likely to form and be stable. However, for a multiprincipal component material, calculating the ΔG_{mix} of all possible phases was impossible. Therefore, a simple method was needed to predict the formation of solid solutions. Based on the equal stoichiometric ratios of RE elements and ignoring the effects of vacancy and nonstoichiometric ratio on configurational

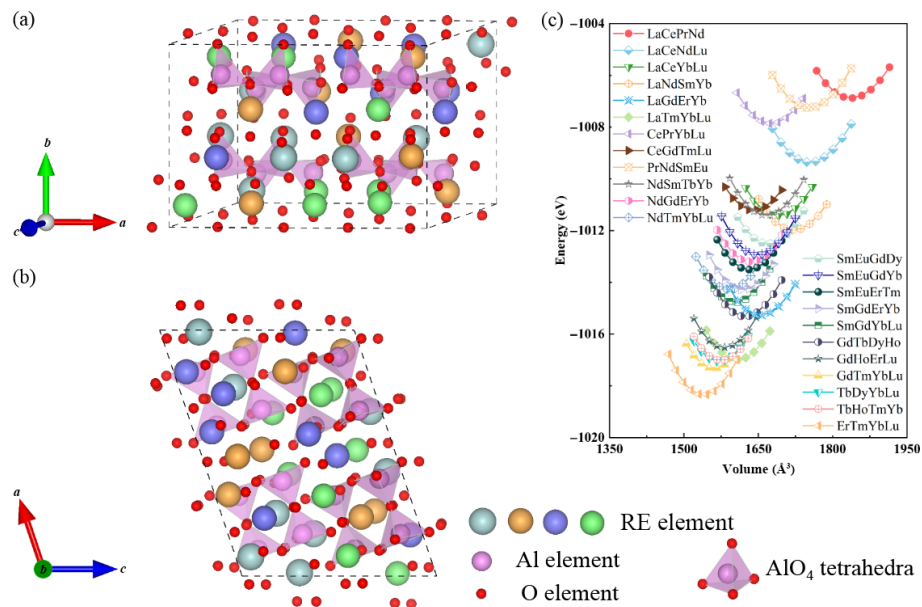


Fig. 1 (a, b) Characterization of crystal structure for $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})\text{Al}_2\text{O}_9$ with a $2 \times 1 \times 1$ supercell. (c) Variation in calculated total energy with volume for $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})\text{Al}_2\text{O}_9$.

entropy, the ΔG_{mix} of the $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})\text{Al}_2\text{O}_9$ system was described as Eq. (2) [35]:

$$\Delta G_{\text{mix}} = \Delta H_{\text{mix}} - T\Delta S_{\text{mix}} \quad (2)$$

where ΔH_{mix} is the mixing enthalpy, ΔS_{mix} is the mixing entropy, and T is the temperature. The mixing enthalpy was determined by comparing the energy of the $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})\text{Al}_2\text{O}_9$ system with those of four single RE aluminates crystallized in the same space group, which was expressed as Eq. (3) [35]:

$$\Delta H_{\text{mix}} = \frac{E_{\text{total}}^{(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4}\text{Al}_{16}\text{O}_{72})}}{120} - \sum_i \frac{E_{\text{total}}^{(\text{RE}^{\text{I}}_{1/4}\text{Al}_8\text{O}_{36})}}{240} \quad (3)$$

where $E_{\text{total}}^{(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4}\text{Al}_{16}\text{O}_{72})}$ and $\sum_i E_{\text{total}}^{(\text{RE}^{\text{I}}_{1/4}\text{Al}_8\text{O}_{36})}$ are the total energies of multiprincipal component RE aluminates and four single RE aluminates, respectively. Moreover, the mixing entropy was calculated as Eq. (4) [35]:

$$\Delta S_{\text{mix}} = -R \left(\frac{X}{X+Y+Z} \sum_i c_i^h \ln c_i^h + \frac{Y}{X+Y+Z} \sum_i c_i^k \ln c_i^k + \frac{Z}{X+Y+Z} \sum_i c_i^l \ln c_i^l \right) \quad (4)$$

where R is the ideal gas constant, and $R = 8.314 \text{ J}/(\text{mol}\cdot\text{K})$. The system contained three different sublattices: h stands for the A-site sublattice with a number of sites X , k stands for the B-site sublattice with a number of sites Y , and l stands for the anionic sublattice with a number of sites Z . Among $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})\text{Al}_2\text{O}_9$, the A-site sublattice is randomly occupied by four different equimolar RE ions, and the B-site sublattice is occupied by aluminum ions, while the anionic sublattice is occupied by oxygen ions. c_i^h , c_i^k , and c_i^l are the contents of h , k , and l , respectively. Based on Boltzmann's hypotheses, the ΔS_{mix} for a regular solution with n components was given by Eq. (5):

$$\Delta S_{\text{mix}} = -R \sum_i c_i \ln c_i \quad (5)$$

For the equiatomic ratio ceramic, ΔS_{mix} can be expressed by Eq. (6):

$$\Delta S_{\text{mix}} = R \ln n \quad (6)$$

where ΔS_{mix} reaches its maximum value.

2.2 Machine learning

Machine learning is a powerful data-driven paradigm that leverages existing data to learn complex structure-performance relationships and make predictions. High-quality data were the cornerstone of effective feature selection and accurate model prediction. With the aim of establishing a reliable training dataset, twenty-seven single-phase aluminates and twenty-seven multiphase aluminates were identified from relevant literature and previous studies [15,36]. During the training process, specimens representing single-phase and multiphase were labeled "1" and "0", respectively. As a foundational procedure in machine learning, feature engineering directly governs the theoretical upper bound of model performance. Accordingly, machine learning models and algorithms were intended to optimize performance as much as possible within these limitations. Within this framework, the mean value ($\overline{\text{prop}}$) and the deviation (σ_{prop}) of properties served as key metrics, expressed as Eqs. (7) and (8):

$$\overline{\text{prop}} = \sum_i c_i \text{prop}_i \quad (7)$$

$$\sigma_{\text{prop}} = \sqrt{\sum_i c_i \left(1 - \frac{\text{prop}_i}{\overline{\text{prop}}} \right)^2} \quad (8)$$

where prop_i denotes the value associated with the i -th constituent $\text{RE}_4\text{Al}_2\text{O}_9$. For the prediction of single-phase formation ability, properties that can be considered include the atomic number, ionic radius, electronegativity, mass, lattice constant, melting point, and density of the corresponding single-principal RE aluminate.

2.3 Experiments

In this study, multiprincipal component RE aluminates ($\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4}\text{Al}_2\text{O}_3$) were synthesized via a two-step experimental strategy including powder synthesis and hot-pressing sintering. The specific procedures were as follows: First, the RE oxide powder of RE_2O_3 (RE = Nd, Sm, and Eu) was calcined at 1300 °C for 2 h to remove residual moisture and organic impurities in the oxides. Next, high-purity RE_2O_3 (RE = Nd, Sm, Eu, Gd, Dy, Ho, Er, Yb, and Lu) and Al_2O_3 powder (Table S1 in the Electronic Supplementary Material (ESM)) were weighed according to the stoichiometric ratio ($\text{RE}_{\text{total}} : \text{Al} = 2 : 1$), ground at 250 r/min for 24 h with anhydrous ethanol as the medium, and dried at 80 °C for 12 h to obtain uniformly mixed presintered powder. After screening with a 200-mesh sieve, high-entropy ceramic powder was composited by sintering at 1600 °C for 4 h. Subsequently, the powder was ball-milled again for 48 h and dried. The high-entropy ceramic powder was heated to 1750 °C for 1 h in a graphite mold with a diameter of 50 mm at a biaxial pressure of 30 MPa and a heating rate of 8 °C/min under an argon atmosphere. Eventually, dense ($\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4}\text{Al}_2\text{O}_3$) ceramic specimens were obtained after natural cooling to room temperature.

2.4 Characterization

The phase structures of the multiprincipal RE aluminate ceramics were determined by an X-ray diffractometer (XRD; SmartLab, Rigaku, Japan) with Cu K α radiation at a scanning rate of 5 (°)/min. The microstructure and elemental distribution were characterized using a field emission scanning electron microscope (SEM; GeminiSEM300, Zeiss, Germany) equipped with an energy dispersive X-ray spectrometer (EDS). The elastic moduli at room temperature, including Young's modulus (E) and shear modulus (G), were measured by Impulse Excitation Tester (RFDA-HTVP 1750C, IMCE, Diepenbeek, Belgium) with specimen dimensions of 15 mm $\times$ 3 mm $\times$ 40 mm. Bulk modulus (B) was calculated based on the relationship between E and G through Eq. (9):

$$B = \frac{GE}{9G - 3E} \quad (9)$$

Then, Poisson's ratio (ν) was calculated by B and G , as shown in Eq. (10):

$$\nu = \frac{3B - 2G}{2(3B + G)} \quad (10)$$

The hardness (H_v) of the specimen was tested by a Vickers hardness tester (HXD-1000TMC/LCD, Shanghai Taiming Optical Instrument Co., Ltd., China), as shown in Eq. (11):

$$H_v = 0.464 \times \frac{P}{a^2} \quad (11)$$

where P is the indentation load applied to the indenter, and a is the arithmetic mean of the two diagonals of the indentation. The applied loads were 0.1, 0.49, 0.98, 2.94, 4.9, and 9.8 N, and the dwell time was 10 s. For each load, the average Vickers hardness was calculated from 10 valid indentations, and all specimens were meticulously polished prior to testing to minimize surface artifacts. The flexural strength (σ_f) test was conducted under the conditions of a span of 30 mm and a loading rate of 0.5 mm/min, as shown in Eq. (12):

$$\sigma_f = \frac{3FL}{2bd^2} \quad (12)$$

where F is the applied load, L is the span length, and b and d are

the width and thickness of the specimen, respectively. The fracture toughness (K_{IC}) was estimated by the Vickers indentation method with Eq. (13) [37]:

$$K_{IC} = 0.016 \left(\frac{E}{H_v} \right)^{1/2} \frac{P}{c^{3/2}} \quad (13)$$

where c is the radial crack length measured from the center of the indentation. A monotonic load was applied to the specimen with dimensions of 3 mm $\times$ 4 mm $\times$ 36 mm using a universal testing machine (Instron-5566, Instron, USA). Damage tolerance (D_t) refers to the ability of a material to withstand loads without catastrophic failure in the presence of cracks or damage [38]. It was defined as Eq. (14):

$$D_t = \frac{K_{IC}E}{\sigma_f H_v} \quad (14)$$

The thermal diffusivity (D_{th}) of the specimen was tested by a thermal analyzer (Netzsch LFA 457, Bavaria, Germany) using the laser flash method with a size of $\phi 12.7$ mm $\times$ 2 mm. Based on the oxides (RE_2O_3 and Al_2O_3) composed of the specimen, the specific heat capacity (c_p) was calculated using the thermodynamic database in accordance with the Neumann-Kopp rule [39]. The density (ρ) of the bulk specimen was measured using the Archimedes drainage method. The thermal conductivity (κ) was obtained from the relationship between D_{th} , c_p , and ρ , as shown in Eq. (15):

$$\kappa = D_{th}c_p\rho \quad (15)$$

3 Results

3.1 Thermodynamic stability

The total energy of ($\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4}\text{Al}_2\text{O}_3$) calculated by first-principles calculations was fitted as a function of volume into the third-order Birch-Murnaghan EOS, as illustrated in Fig. 1(c). The symbols represent the original data points calculated, and the solid lines represent the fitted EOS. The energy curve is clearly parabolic, with a distinct minimum total energy at the optimal volume. The specific values for the minimum energy and corresponding equilibrium volume are summarized in Table S2 in the ESM. The variation in equilibrium volume can be attributed to the differing ionic radii of the multiprincipal component RE aluminates, with the overall trend showing a gradual decrease in equilibrium volume as the atomic number increases. Our calculation results align with previous studies [40], validating the reliability of the calculation method adopted in this study. The phase formation ability of high-entropy materials has been intensively studied. Although the presence of multiple principal elements can effectively tailor material properties, chemical disorder also complicates the prediction of phase stability. The ΔG_{mix} as a function of temperature is shown in Fig. S1 in the ESM. According to Eq. (2), a high ΔS_{mix} can significantly reduce ΔG_{mix} , thereby lowering the tendency of ordering and segregation in ceramic materials and promoting the formation of solid solutions.

3.2 Deviation parameters

In structures containing multiple principal elements or complex compositions, the chemical disorder induced by the mixture of multiple cations can optimize various properties of the material. This disorder can be quantified by ionic radius deviation, electronegativity deviation, mass deviation, and so on. Excluding the radioactive element Pm, the remaining fourteen RE elements

can form $C_{14}^4 = 1001$ combinations of four principal high-entropy aluminates. The relationships between the ionic radius deviation and the average ionic radius, the electronegativity deviation and the average electronegativity, as well as the mass deviation and the average mass for the 1001 combinations are plotted in Fig. 2. Nonexperimental combinations, multiphase structures, and single-phase structures are denoted by hollow circles, black hollow shapes with crosses, and five-pointed stars, respectively. Each data point corresponds to a four-principal equal-mole combination. Moreover, 25 multiphase structures and 14 single-phase structures are identified in the experimental combination. As evidenced by the star symbols, single-phase solid solutions are predominantly observed when the ionic radii fall within 0.89–0.94 Å and the ionic radius deviations lie between 2% and 6%. While these ranges provide a useful reference for predicting solid solution formation, they are not intended as an absolute criterion. Building on the established association between these deviation parameters and phase stability, these feature parameters are subsequently utilized as feature descriptors for machine learning models designed to predict and screen novel high-entropy ceramic combinations.

3.3 Feature selection

The appropriate selection of input features is a vital factor in developing supervised machine learning models for research materials. After careful investigation of the elemental attributes and material characteristics of high-entropy RE aluminates, low-cost input features are prioritized for design based on domain knowledge and literature guidance [25]. The input features that do not require complex modeling computations and their corresponding feature descriptors are listed in Table S3 in the ESM. Some selected input features may exhibit a high degree of correlation, which can cause the machine learning model to be more sensitive to highly correlated features and less sensitive to low-correlated features. The Pearson correlation coefficient (r), which assesses the linear relationship between two variables, is selected as a key indicator of prediction quality, with r defined as Eq. (16):

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (16)$$

where n is the number of variables, x and y are two independent input features, and $\bar{x}$ and $\bar{y}$ are their corresponding mean values. The value of r ranges from -1 to $+1$, where -1 and $+1$ represent positive and negative correlations, respectively, and 0 represents no correlation between features. Features with high linear

correlation contain repetitive or similar information, which reduces the learning efficiency. To avoid model redundancy and improve the accuracy of model prediction, features with r exceeding 0.90 are removed. This threshold is adequately high to ensure that the retained features preserve useful information while avoiding excessive feature loss [41,42]. The strong positive and negative correlation features are represented by red and blue (Fig. 3), respectively. The highest correlation is found between the average mass ($\bar{m}$) and average ionic radius ($\bar{r}$), which is -0.89 . Moreover, the absolute values of r for average electronegativity ($\bar{\chi}$) and $\bar{r}$, electronegativity deviation (σ_χ) and RE ionic radius deviation (σ_r), as well as $\bar{m}$ and $\bar{\chi}$, are all greater than 0.5 , which indicates a strong correlation among them. The minimum eigenvalue of the correlation matrix is 0.03 , indicating that there is a reasonable linear independence between the features. The data in the figure confirm that no overfitting occurred. Therefore, all seven input features selected during the training process are reasonable.

The selected material feature descriptors vary in terms of magnitude, unit, and scope. Since most machine learning algorithms use only the Euclidean distance between two data points, material descriptors with higher magnitudes have greater weight than those with lower magnitudes. To realize the same level of magnitude for all feature descriptors and accelerate the convergence speed of the machine learning model, each material descriptor is normalized through Eq. (17) [43]:

$$x_i^{\text{norm}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (17)$$

where $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the feature descriptor in the training model, respectively.

3.4 Hyperparameter optimization

The purpose of hyperparameter optimization is to obtain an improved machine learning model with high prediction accuracy and universality. Three machine learning models, namely, artificial neural network (ANN), support vector machine (SVM), and random forest classification (RFC), are utilized to predict the single-phase formation ability of materials in this study. The ultimate goal is to reliably identify four principal combinations with stable phase structures. This consensus-based screening strategy minimizes the risk of overreliance on any single algorithmic bias and enhances the reliability of the final results. The key structural parameters of the ANN model, including the number of hidden layers and nodes per layer, are optimized through a systematic grid search. Insufficient hidden layers or nodes per layer limit feature extraction, while excessive complexity can introduce noise and overfitting risks [44]. Therefore, it is necessary to strike a balance between model capacity and

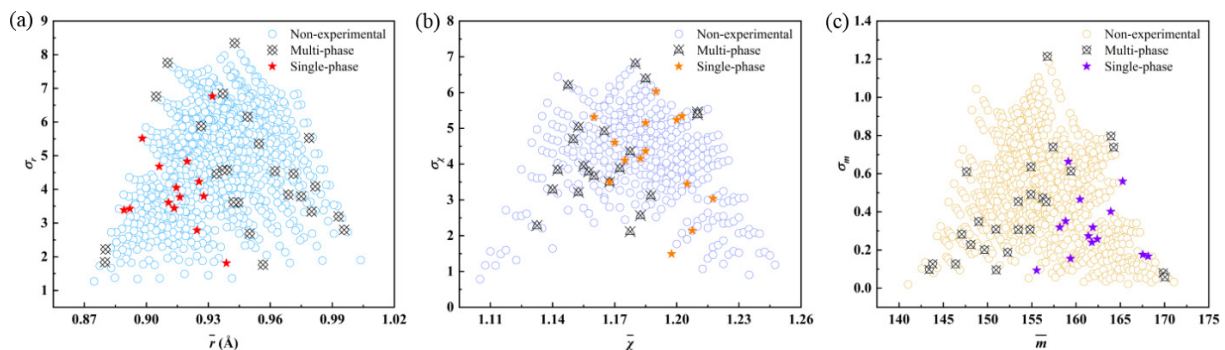


Fig. 2 Disorder parameters of $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})\text{Al}_2\text{O}_9$ ceramics: (a) ionic radius deviation, (b) electronegativity deviation, and (c) mass deviation.

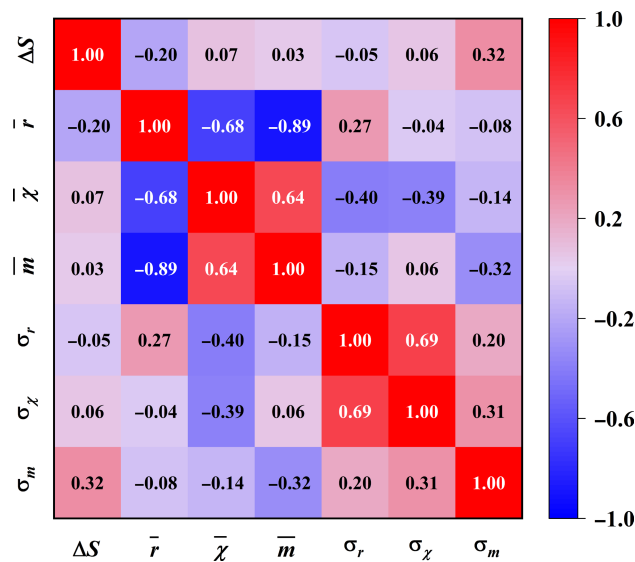


Fig. 3 Plot of Pearson correlation coefficients between each pair of input features.

generalization ability, with the focus on obtaining the minimum network configuration that can maintain strong predictive performance on the validation data. The optimization of the SVM model involves the refined adjustment of critical hyperparameters, particularly the kernel function and regularization parameter (C). The kernel function enables linear separability by mapping input data into high-dimensional feature spaces, and C governs the trade-off between maximizing the margin and minimizing classification errors [45,46]. Higher C values produce narrower marginal hyperplanes, which more strictly penalize misclassified training samples. The RFC model is an ensemble prediction tool based on classification and regression tree analysis [47,48]. The core is to integrate the prediction results of multiple decision trees to improve the overall prediction accuracy while mitigating overfitting. The ANN excels at capturing intricate hierarchical relationships, and the SVM model provides strong theoretical guarantees for generalization, while the RFC model delivers reliable predictions with inherent feature importance measures. These machine learning methods exhibit complementary advantages, collectively offering comprehensive solutions to diverse classification challenges.

The training set and test set are randomly divided at ratios of 0.85:0.15, 0.8:0.2, and 0.75:0.25 using the random splitting technique. The accuracy and rationality of each model are evaluated by measuring the prediction accuracy at three ratios. The respective validation accuracies achieved by the three machine learning models at different ratios are shown in Fig. 4. The ANN model attains optimal validation accuracy (1.00) at test set ratios of 0.15 and 0.2, demonstrating strong generalization performance. However, when the test set ratio increases to 0.25, the validation accuracy of the ANN model significantly declines to 0.57, indicating reduced robustness under a larger test partition. The SVM model exhibits similar validation accuracy across different test set ratios, which can reveal consistent but suboptimal generalization performance. Although this stability reflects that the SVM model has a certain degree of robustness in dataset partitioning, the suboptimal validation accuracies imply the limitations of the model's learning capacity, which may be caused by kernel mismatch or inadequate feature separation. The RFC model illustrates the optimal verification accuracy (1.00) under three test set ratios, proving robust generalization performance

regardless of dataset partitioning. This stability indicates that the model is relatively insensitive to variations in the training-test split, likely due to its effective margin maximization and regularization mechanisms. In short, both the ANN and RFC models achieve optimal validation accuracy, representing their outstanding ability to capture complex patterns from the given dataset. The performance advantage may be attributed to the hierarchical feature learning mechanism of ANN and the integrated decision strategy of RFC, which significantly enhance the generalization ability of the models.

3.5 Model evaluation

The evaluation of machine learning models typically relies on three fundamental tools: the confusion matrix, receiver operating characteristic (ROC) curve, and area under the curve (AUC) score, which provide comprehensive insights into model performance across classification thresholds. In binary classification, the four modules of the confusion matrix are true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP and TN reflect the model's ability to correctly classify, with higher values indicating better classification performance. The error types FP and FN offer critical diagnostic information about the model, with their relative importance being highly dependent on the actual situation. As shown in Figs. 5(a) and 5(b), the confusion matrices of the RFC and ANN models exhibit similar patterns. The RFC and ANN models achieved perfect classification performance on the current test data. The outstanding classification quality of the RFC and ANN models is conclusively demonstrated by their perfect AUC scores of 1.0, as derived from the ROC curves (Fig. 5(c)). The empirical observation of near-1 AUC values indicates that the predictions are of reasonable quality [49]. Although the ROC curve provides a powerful description of the model's performance, it may overestimate the model's effectiveness on small datasets. High overall accuracy does not imply consistent prediction for all individual samples. Instead, the observed disagreements near the decision boundary underscore the generalization constraints imposed by the limited sample size. The precision-recall (PR) curve is further utilized to evaluate model performance, as depicted in Fig. 5(d). The closer the PR curve is to the upper right corner, the better the model performance. The analysis of ROC and PR yields convergent evidence, making our conclusion more convincing. However, such excellent results warrant further investigation to avoid potential data simplification and overfitting.

To improve the reliability and transparency of the results and gain a deeper understanding of the contribution of each feature to the model prediction, SHapley Additive exPlanations (SHAP) analysis is introduced to enhance the interpretability of machine learning models [50–52]. The fundamental principle of SHAP is to quantitatively assess the extent to which each feature affects the final model output [53]. Specifically, this method quantifies the relationship between the inputs and outputs of the model, providing insights into “black box” models from both global and local perspectives. The SHAP analysis for the RFC and ANN models elucidates the relative importance of each feature in predicting the single-phase formation ability of $(\text{RE}^{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics, as illustrated in Figs. 6(a) and 6(c). Features are vertically arranged in descending order according to their average $|\text{SHAP}|$ values, and the x -axis represents the corresponding $|\text{SHAP}|$ values. For both models, the $\bar{r}$ of materials ranks first in predicting the single-phase formation ability. The differences in $\bar{r}$ of RE elements within the material may lead to more lattice distortions and defects. Moreover, we can observe that there are some differences in the importance of other

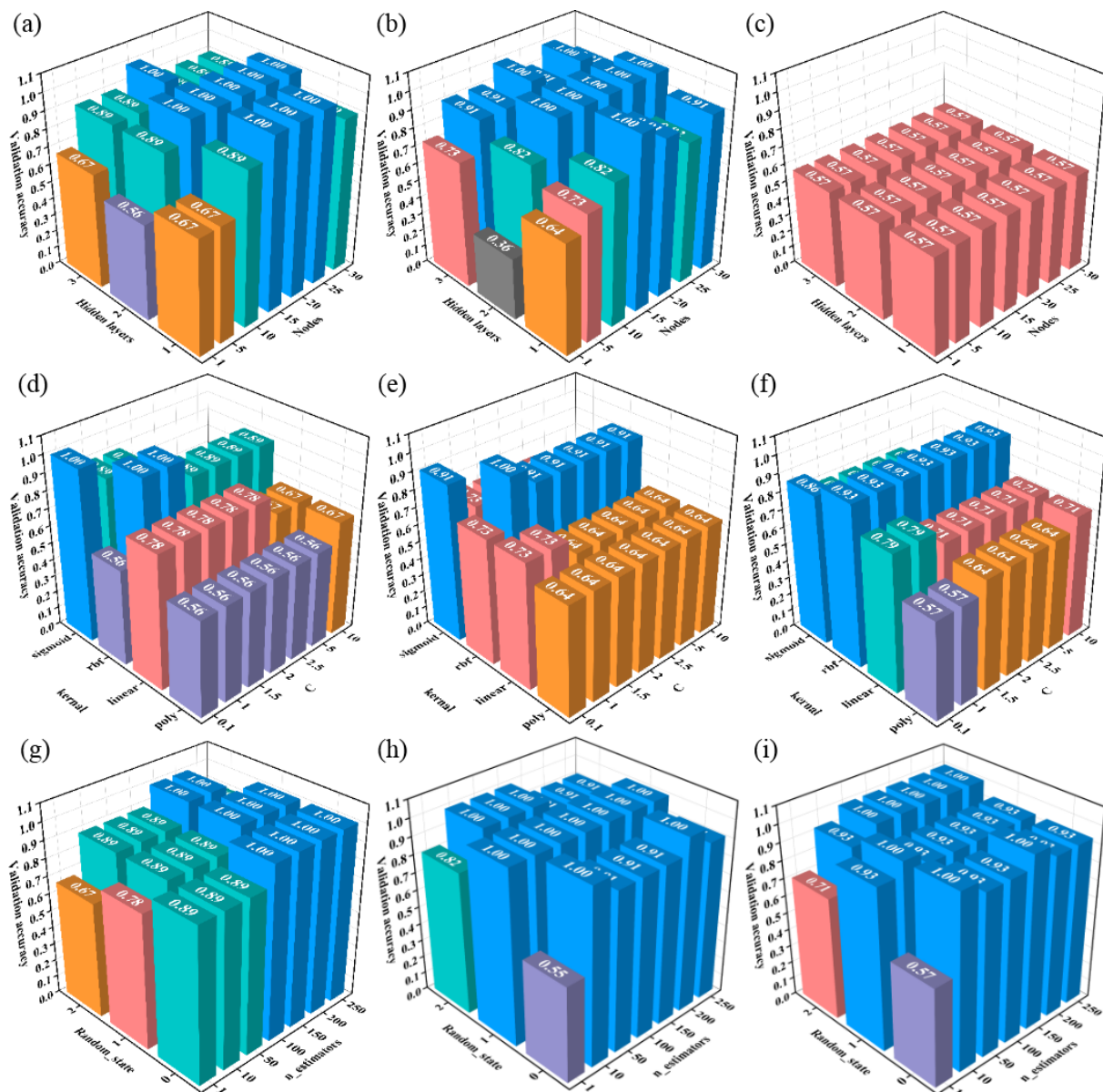


Fig. 4 Hyperparameter optimization of (a–c) ANN, (d–f) SVM, and (g–i) RFC models. Test ratios for the first, second, and third columns are 0.15, 0.2, and 0.25, respectively.

features to the outputs of the two models. For the RFC model, $\bar{r}$ plays an absolutely dominant role, followed by $\bar{\chi}$, $\bar{m}$, and ΔS , while σ_r , σ_χ , and σ_m have comparatively minor impacts. For the ANN model, $\bar{m}$ demonstrates significant contributions alongside $\bar{r}$. Furthermore, the effects of σ_χ and σ_m are similarly small in both models. Individual data points in the scatter plot denote samples (Figs. 6(b) and 6(d)), with red and blue hues indicating higher and lower feature values, respectively. A positive SHAP value implies that the feature enhances the predicted target attribute, while a negative SHAP value indicates that its effect is weakened [54]. The visualization graph intuitively and clearly demonstrates how each feature affects the prediction result. The influence and importance of features on the model output are directly proportional to their horizontal coverage. In other words, higher-ranked features display more widely distributed SHAP values, although their dependence on feature values follows nonlinear patterns, especially for features occupying lower ranks [55]. The distribution range of SHAP values for $\bar{r}$ is significantly wider than that of other features, indicating that $\bar{r}$ has the greatest

impact on the prediction of the RFC and ANN models. The $\bar{r}$ is directly linked to the lattice distortion and polyhedral connectivity of the RE–O frameworks. The coexistence of positive and negative SHAP contributions reveals a size-driven nonmonotonic stabilization mechanism, indicating the existence of an optimal $\bar{r}$ range for the stable $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})\text{Al}_2\text{O}_9$ phase.

3.6 Prediction result

The well-trained RFC and ANN models are employed to assess the single-phase formation ability of $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})\text{Al}_2\text{O}_9$ candidates. For the purposes of experimental validation, the expensive RE elements Pr and Tb are excluded. Furthermore, the RE element Gd is chosen for its ability to offer the balance of ionic radius, electronegativity, and atomic mass among the four principal elements. Therefore, this yields a total of $C_{11} = 165$ four-component combinations. Excluding 23 combinations of existing experimental data, the remaining 142 combinations are applied to the subsequent single-phase formation capability assessment via machine learning. To the best

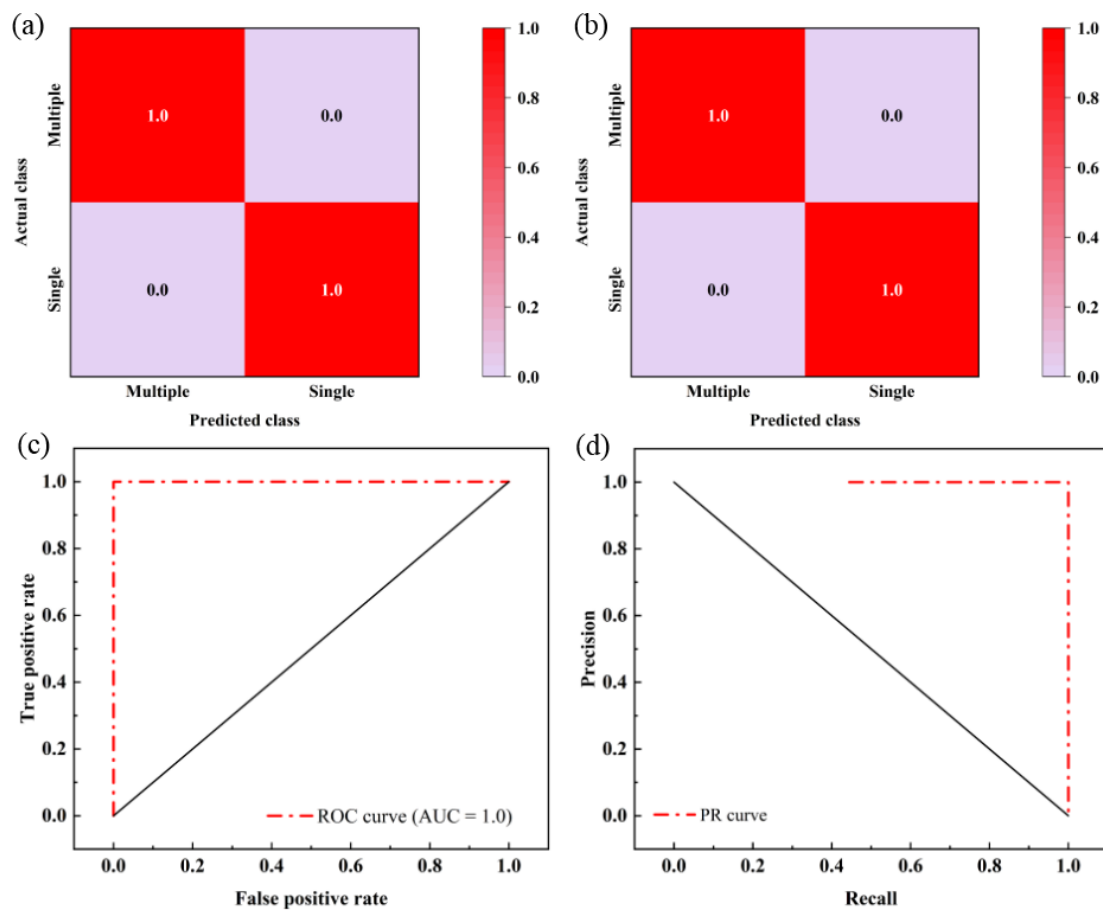


Fig. 5 Performance of RFC and ANN models. (a, b) Confusion matrix, (c) ROC-AUC curve, and (d) PR curve.

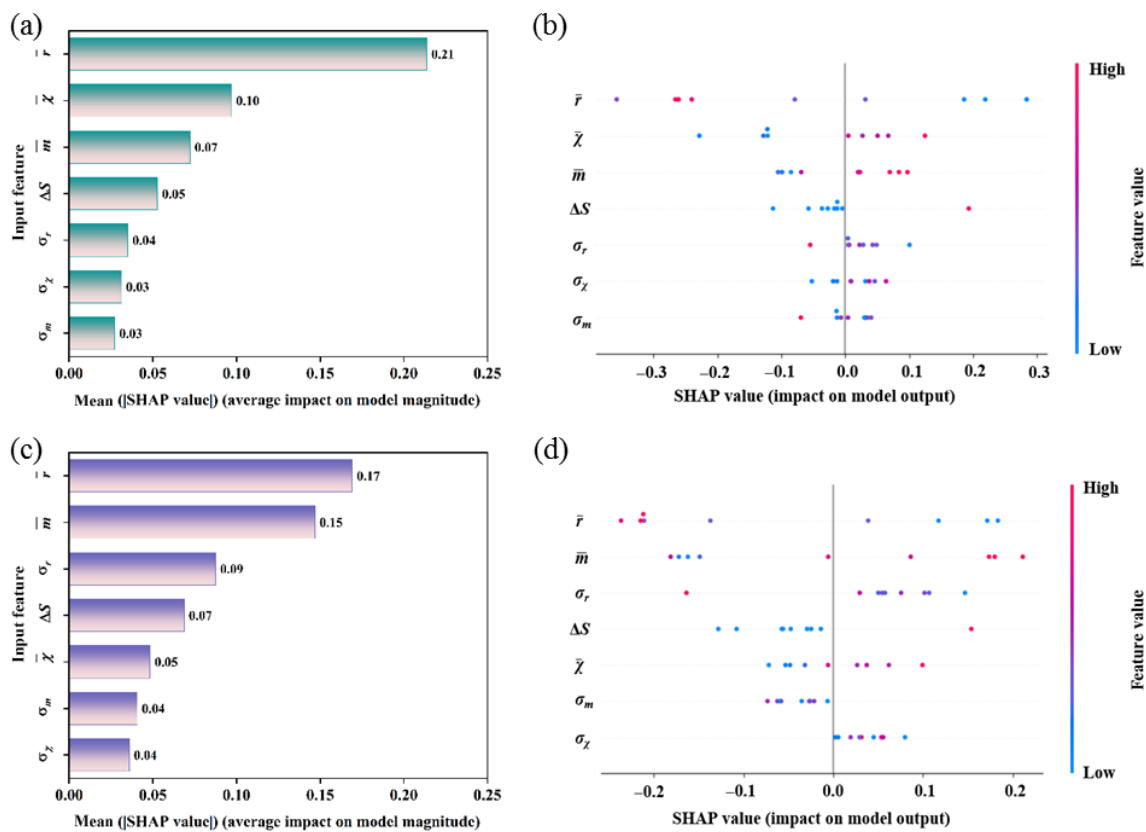


Fig. 6 Feature importance of (a) RFC and (c) ANN models. SHAP values for each feature in (b) RFC and (d) ANN models.

of our knowledge, none of these candidate materials have been experimentally synthesized or reported before. The prediction results for single-phase formation ability are summarized in Table S4 in the ESM. Among the 142 predicted samples, 81 samples are classified as pure $(\text{RE}^{I}_{1/4}\text{RE}^{II}_{1/4}\text{RE}^{III}_{1/4}\text{RE}^{IV}_{1/4})_4\text{Al}_2\text{O}_9$ by the RFC model, while the ANN model identifies 70 single phases. A slight difference is observed between the RFC and ANN models, with only 13 samples of discordant prediction results. Figure S2 in the ESM shows the XRD pattern with inconsistent prediction results between the RFC and ANN models. To predict phase stability more accurately, this study aims to identify consistent conclusions by comparing the prediction results of different machine learning models to reduce the possible bias caused by relying on a single model. Consequently, three samples of four principal combinations are randomly selected from 69 single-phase results of the RFC and ANN models with consistent predictions for experimental verification, with the aim of confirming the reliability of the machine learning results.

To assess the accuracy of the RFC and ANN models, a leave-one-out cross-validation (LOOCV) approach is adopted, with ROC-AUC curves depicted in Fig. S3 in the ESM. The LOOCV results yield an accuracy of 0.76 and an AUC of 0.91, confirming that this approach serves as a stable and reliable evaluation strategy. The decrease in AUC reflects the inherent stringency of the LOOCV approach, which can evaluate the model on every sample individually, including boundary or hard-to-classify cases that are unlikely to be concentrated together in a single small test set from a random split. The LOOCV validation results for the predicted phase formation ability are provided in Table S4 in the ESM. The high consistency between the LOOCV predictions and those from the RFC and ANN models confirms that the latter have reliably captured the intrinsic compositional features governing phase formation. The convergence of the RFC and ANN models is further corroborated by LOOCV-based validation. The consistency confirmed by this most stringent cross-validation approach strongly attests to the robustness of the predicted phase formation outcomes. Notably, the LOOCV approach effectively rules out the possibility that the results of this study are merely false positives caused by overfitting or favorable data partitioning, thereby significantly enhancing the credibility of the samples subsequently selected for experimental validation.

3.7 Experiments

According to the distinction between light and heavy RE elements,

the predicted combinations of the four principal high-entropy aluminates can be divided into four categories: three light rare earth elements (LREEs) with one heavy rare earth element (HREE), two LREEs with two HREEs, one LREE with three HREEs, and four HREEs. As indicated by the prediction results, four principal systems comprising one LREE with three HREEs fail to form single-phase structures. As a result, these combinations are omitted from subsequent experimental validation. On the basis of their reactivity toward CMAS corrosion, the RE elements investigated can be classified into three groups: active elements (Gd, Tb, and Dy), neutral elements (Ho, Er, and Tm), and inert elements (Yb and Lu) [25]. High-entropy RE aluminates $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ are selected for experimental validation. The three selected combinations cover a broad range of the most important input features, with $\bar{r}$ being 0.939, 0.920, and 0.898 Å, respectively (Table S5 in the ESM). This wide coverage across the feature space ensures that the experimental validation provides a robust test of the predictive capability of RFC and ANN models. To ascertain the phase formation ability of the synthesized high-entropy ceramics, XRD analyses were performed on the samples. The XRD patterns of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$, along with the PDF cards for their respective single RE aluminates, are illustrated in Fig. 7. The three synthesized ceramics are characterized as having a monoclinic crystal structure with a $P2_1/c$ space group. In comparison to the PDF cards of the four single RE aluminates, the diffraction peaks of the high-entropy ceramic exhibit a slight shift, indicating that the multielement mixture at the A-site leads to lattice distortion. The single-phase structures of three high-entropy RE aluminates are determined by XRD analyses, which are consistent with the prediction results of machine learning.

SEM-EDS analysis provides further evidence for the successful formation of single-phase high-entropy aluminates. The SEM images and mappings of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics are depicted in Figs. 8(a)–8(c). The three high-entropy ceramic materials display relatively dense microstructures with lower porosity. The EDS mapping confirms that the sample consists entirely of the expected elements, with no impurity elements detected. In addition, all elements are uniformly distributed across the scanned area, which can indicate that solid solutions have been successfully formed. The elemental

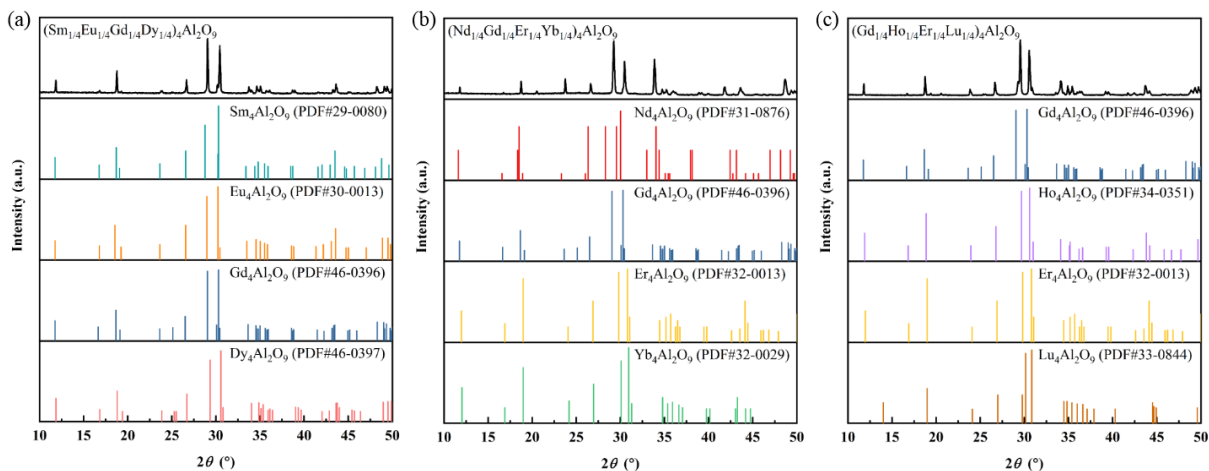


Fig. 7 XRD patterns of (a) $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, (b) $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and (c) $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$, along with PDF cards for their respective single RE aluminates.

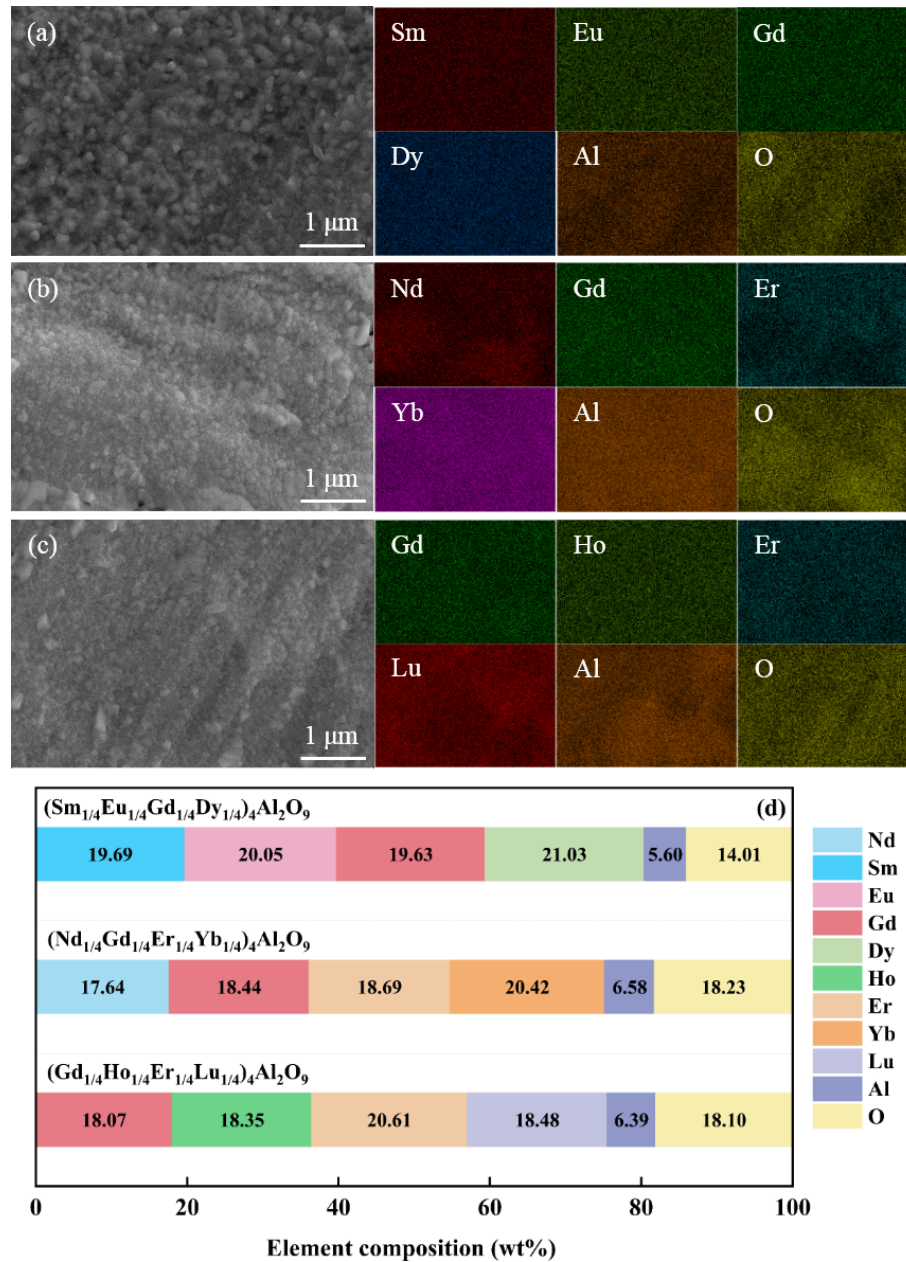


Fig. 8 SEM images and mappings of (a) $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, (b) $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and (c) $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics. (d) Corresponding elemental composition obtained by EDS (wt%).

Table 1 Young's modulus (GPa), shear modulus (GPa), bulk modulus (GPa), Pugh's ratio, Poisson's ratio, Vickers hardness (GPa, at 9.80 N), flexural strength (MPa), fracture toughness ($\text{MPa}\cdot\text{m}^{1/2}$), and damage tolerance ($\text{m}^{1/2}$) of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$, and other reference ceramics [13,56]

	$(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$	$(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$	$(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$	$(\text{La}_{0.25}\text{Sm}_{0.25}\text{Gd}_{0.25}\text{Ho}_{0.25})\text{AlO}_3$	Al_2O_3
E (GPa)	170.4	218.4	205.5	283.96	403
G (GPa)	67.1	86.6	81.5	114.63	163
B (GPa)	123.3	152.3	143.1	181.05	254
G/B	0.544	0.569	0.570	0.63	0.642
ν	0.27	0.26	0.26	0.24	0.23
H_v (GPa)	7.72 ± 0.15	9.20 ± 0.13	8.88 ± 0.12	15.90 ± 0.55	14.4
σ_f (MPa)	170.9	216.5	195.9	150.19	379
K_{IC} ($\text{MPa}\cdot\text{m}^{1/2}$)	2.5 ± 0.07	3.2 ± 0.16	2.9 ± 0.11	2.49 ± 0.08	4
D_t ($\text{m}^{1/2}$)	0.32	0.35	0.34	0.30	0.27

composition of the high-entropy ceramics was quantified by EDS semiquantitative analysis. As shown in Fig. 8(d), the measured elemental ratios are in good agreement with the theoretical stoichiometry, with only minor deviations. The successful synthesis of single-phase solid solutions is confirmed by XRD phase verification and SEM morphological analysis, which is highly consistent with the prediction of machine learning. The consistency between experimental characterization and machine learning screening emphasizes the predictive capability of the selected machine learning models in identifying the stability of high-entropy phases.

The mechanical response of $(\text{RE}^{I}_{1/4}\text{RE}^{II}_{1/4}\text{RE}^{III}_{1/4}\text{RE}^{IV}_{1/4})_4\text{Al}_2\text{O}_9$ materials is strongly dependent on their crystalline phases and microstructural uniformity. Therefore, it is necessary to establish a link between microstructural characteristics and mechanical properties. The detailed mechanical properties of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ are compiled in Table 1, including elastic constants, hardness, and fracture-related parameters. The shear modulus of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ is lower than the bulk modulus, which can make them relatively prone to shear-related structural deformation and help to increase the damage tolerance of the materials. In contrast to previously reported $(\text{La}_{0.25}\text{Sm}_{0.25}\text{Gd}_{0.25}\text{Ho}_{0.25})\text{AlO}_3$ and Al_2O_3 [13], the three ceramics exhibit lower B , G , and E , which is in agreement with DFT predictions [40]. Pugh's ratio (G/B) is widely used to distinguish the brittleness and quasiductility of materials. The G/B values of

$(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ are determined to be 0.544, 0.569, and 0.570, respectively, all of which are below the critical threshold of 0.571. Therefore, all three RE aluminates are classified as quasiduct ceramics. Unlike Al_2O_3 , which contains only strong and rigid Al–O bonds, $(\text{RE}^{I}_{1/4}\text{RE}^{II}_{1/4}\text{RE}^{III}_{1/4}\text{RE}^{IV}_{1/4})_4\text{Al}_2\text{O}_9$ exhibits relatively weaker RE–O bonds and multiple RE^{3+} ions. Lattice mismatch among RE^{3+} ions induces local distortion, which impedes dislocation slip and contributes to the suppression of microcrack initiation. Concurrently, the sluggish diffusion effect retards grain growth and slows crack propagation. Finally, these effects enhance the damage tolerance of materials. Moreover, unlike the rigid three-dimensional network framework formed by AlO_6 octahedra in $(\text{La}_{0.25}\text{Sm}_{0.25}\text{Gd}_{0.25}\text{Ho}_{0.25})\text{AlO}_3$, a layered/chain-like structure facilitating stress dissipation is formed in $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$, accounting for the quasiductility of the studied material.

Furthermore, Poisson's ratio, Vickers hardness, flexural strength, fracture toughness, and damage tolerance are also important indicators for evaluating the mechanical properties of ceramics. The $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics exhibit higher Poisson's ratios ($\nu > 0.25$) than those of $(\text{La}_{0.25}\text{Sm}_{0.25}\text{Gd}_{0.25}\text{Ho}_{0.25})\text{AlO}_3$ and Al_2O_3 , implying a lower Pugh's ratio and better ductility. As shown in Fig. 9(a), the Vickers hardness of the $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics varies with the applied loads.

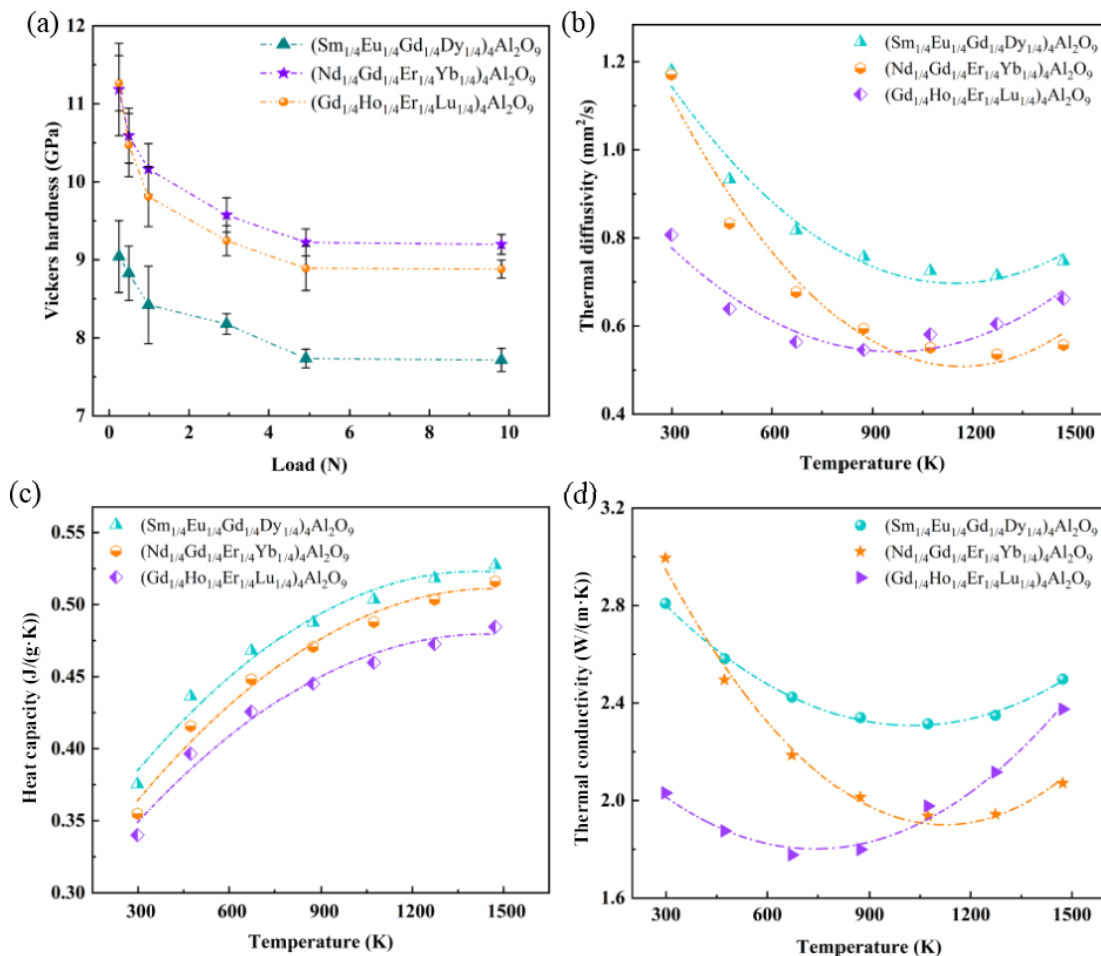


Fig. 9 (a) Variation in Vickers hardness under different applied loads. Temperature-dependent (b) thermal diffusivity, (c) specific heat capacity, and (d) thermal conductivity of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ and other reference ceramics [57].

The Vickers hardness gradually decreases with increasing applied load until a stable value is reached. As the applied load increases from 0.2452 to 9.80 N, the H_v of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$ gradually decreases from 9.04 ± 0.46 to 7.72 ± 0.15 GPa; the H_v of $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$ gradually decreases from 11.18 ± 0.59 to 9.2 ± 0.13 GPa; and the H_v of $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ gradually decreases from 11.26 ± 0.35 to 8.88 ± 0.12 GPa. Figure S4 in the ESM displays SEM images of Vickers hardness indentation produced on three high-entropy ceramics under applied loads of 0.2452, 0.49, 0.98, 2.94, 4.90, and 9.80 N. Unlike $(\text{La}_{0.25}\text{Sm}_{0.25}\text{Gd}_{0.25}\text{Ho}_{0.25})\text{AlO}_3$ and Al_2O_3 , these materials possess lower H_v but higher σ_f and K_{IC} , which suggests outstanding damage tolerance and remarkable resistance to thermal shock-induced cracking, making them highly attractive for the applications of T/EBCs. Overall, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$ exhibits higher H_v , σ_f , K_{IC} , and D_I than $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$, which in turn are higher than those of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$. This trend is related to the ionic radius difference parameters at the A-site of the three high-entropy ceramics, which leads to significant lattice distortion. The elastic stress field generated by lattice distortion can effectively hinder dislocation motion, thereby enhancing the mechanical properties of materials.

The lattice distortion that leads to the solid solution strengthening of high-entropy ceramics inevitably affects phonon scattering, thereby playing a regulatory role in thermal properties. Figures 9(b)–9(d) depict the temperature-dependent thermal diffusivity, specific heat capacity, and thermal conductivity of the $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics, respectively. The thermal diffusivities of the three materials exhibit a nonmonotonic temperature dependence. Taking $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$ as an example, the thermal diffusivity decreases from 1.17 to 0.54 W/(m·K) as the temperature rises from 298 to 1273 K and then slightly increases to 0.56 W/(m·K) at 1473 K. Compared with thermal diffusivity, specific heat capacity increases monotonically with the rise in temperature. The C_p of $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics ranges from 0.341–0.375 J/(g·K) (298 K) to 0.485–0.527 J/(g·K) (1473 K). This indicates that the specific heat capacity is primarily determined by the phonon state density associated with lattice vibrations and is independent of the average phonon mean free path. The thermal conductivities of the $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics reach their lowest values of 2.31 W/(m·K) (1073 K), 1.94 W/(m·K) (1273 K), and 1.78 W/(m·K) (673 K), respectively. In addition, their thermal conductivities are 2.5, 2.07, and 2.37 W/(m·K) at 1473 K, respectively. The Umklapp and defect phonon scattering in the low-temperature region increases with increasing temperature, resulting in a decrease in thermal conductivity. This is attributed to the synergistic optimization of differences in ionic radii, mass differences, and lattice distortion. In the high-temperature region, the thermal conductivity increases slightly due to the influence of thermal radiation. The $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics have lower $\kappa_{\min}$ than $\alpha\text{-Al}_2\text{O}_3$ (2.89 W/(m·K)) [56]. The weaker RE–O polyhedra provide “weak regions” that scatter phonons and reduce the phonon mean free path. The measured thermal conductivity is also lower than or equal to that of conventional materials such as 8YSZ (~2.5 W/(m·K)) and ScTaO_4 (2.37 W/(m·K)) [57]. Taken together, the combination of high melting points ($\text{RE}_4\text{Al}_2\text{O}_9$, 1900–2060 °C [40]), moderate damage tolerance, intermediate moduli, and low thermal conductivity highlights their potential as next-generation T/EBCs.

4 Conclusions

In summary, this study proposes an acceleration strategy based on first-principles calculations and machine learning for the discovery of single-phase high-entropy RE aluminates. The variation in equilibrium volume derived from the third-order Birch–Murnaghan EOS can be attributed to the differing ionic radii of the multiprincipal component RE aluminates, with the overall trend showing a gradual decrease in equilibrium volume as the atomic number increases. Three machine learning models (ANN, SVM, and RFC) are developed to systematically evaluate the single-phase formation ability of $(\text{RE}^{\text{I}}_{1/4}\text{RE}^{\text{II}}_{1/4}\text{RE}^{\text{III}}_{1/4}\text{RE}^{\text{IV}}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics. The seven input features selected during the training process are proven to be appropriate, with no evidence of overfitting. Both the ANN and RFC models achieve perfect validation accuracy and AUC scores, demonstrating their remarkable capability to extract complex patterns from the dataset. In contrast, the SVM model exhibits consistent but suboptimal generalization performance. The trained models evaluate 142 unreported four-component combinations, yielding 69 consistently predicted single-phase candidates.

The experimental results confirm that the prepared ceramics possess a single-phase crystalline structure, which is consistent with the machine learning predictions. All three RE aluminates are classified as quasiductile ceramics. The hardness values of the $(\text{Sm}_{1/4}\text{Eu}_{1/4}\text{Gd}_{1/4}\text{Dy}_{1/4})_4\text{Al}_2\text{O}_9$, $(\text{Nd}_{1/4}\text{Gd}_{1/4}\text{Er}_{1/4}\text{Yb}_{1/4})_4\text{Al}_2\text{O}_9$, and $(\text{Gd}_{1/4}\text{Ho}_{1/4}\text{Er}_{1/4}\text{Lu}_{1/4})_4\text{Al}_2\text{O}_9$ ceramics under an applied load of 9.8 N are 7.72 ± 0.15 , 9.2 ± 0.13 , and 8.88 ± 0.12 GPa, respectively. The three samples exhibit a temperature-dependent thermal conductivity that first decreases and then increases with temperature, reaching minimum values of 2.31, 1.94, and 1.78 W/(m·K) at 1073, 1273, and 673 K, respectively. The excellent consistency between machine learning predictions and experiments accelerates the discovery of high-entropy ceramics, paving the way for rational design and effective property tuning within the immense compositional space.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 52172071), the Outstanding Youth Fund of Henan Province (No. 252300421006), and Shanghai Technical Service Center for Advanced Ceramics Structure Design and Precision Manufacturing (No. 20DZ2294000).

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article. The author Bin Liu is the Editorial Committee member of this journal.

Electronic Supplementary Material

Supplementary material is available in the online version of this article at <https://doi.org/10.26599/JAC.2026.9221354>.

References

- [1] Rost CM, Sachet E, Borman T, et al. Entropy-stabilized oxides. *Nat Commun* 2015, 6: 8485.
- [2] Miracle DB, Senkov ON. A critical review of high entropy alloys and related concepts. *Acta Mater* 2017, 122: 448–511.

- [3] Nie XD, Xu JY, Zhang ZC, *et al.* The progress of research on high-entropy ceramics in the field of tribology. *Adv Ceram* 2025, **46**: 1–23. (in Chinese)
- [4] Zhao ZF, Ruan ZY, Li R, *et al.* High entropy pyrochlore $(\text{La}_{0.3}\text{Gd}_{0.3}\text{Ca}_{0.4})_2(\text{Ti}_{0.2}\text{Zr}_{0.2}\text{Hf}_{0.2}\text{Nb}_{0.2}\text{Ta}_{0.2})\text{O}_7$ ceramic with amorphous-like thermal conductivity for environmental/thermal barrier coating applications. *J Mater Sci Technol* 2025, **205**: 315–326.
- [5] Liu BS, Zhao JL, Chu KL, *et al.* Unraveling the atomic-scale mechanisms of hydrogen defects behavior in yttria-stabilized tetragonal zirconia by first principles calculation. *J Adv Ceram* 2025, **14**: 9221099.
- [6] Li Y, Zhang MY, Chen JN, *et al.* Ceramic-based electromagnetic interference shielding materials: Mechanisms, optimization strategies, and pathways to next-generation applications. *J Adv Ceram* 2025, **14**: 9221194.
- [7] Cui N, Zhang YX, Wang LJ, *et al.* Single-phase formation process and carbon vacancy regulation of $(\text{TiVNbMoW})\text{C}_x$ high-entropy ceramics. *J Inorg Mater* 2025, **40**: 511. (in Chinese)
- [8] Wang YL, Zhong CQ, Zhang JZ, *et al.* Machine learning for predictive design and optimization of high-performance thermoelectric materials: A review. *J Mater Inform* 2025, **5**: 41.
- [9] Divilov S, Eckert H, Hicks D, *et al.* Disordered enthalpy-entropy descriptor for high-entropy ceramics discovery. *Nature* 2024, **625**: 66–73.
- [10] Liu DR, He WT, Wei LL, *et al.* Enhanced radiative cooling of columnar thermal barrier coatings at ultrahigh temperatures and mechanisms underneath. *J Mater Sci Technol* 2025, **239**: 81–92.
- [11] Padture NP. Advanced structural ceramics in aerospace propulsion. *Nat Mater* 2016, **15**: 804–809.
- [12] Li YR, Pan DH, Cao JH, *et al.* Recent advances in high-entropy ceramics: Design principles, structural characteristics, and emerging properties. *Extreme Mater* 2025, **1**: 42–72.
- [13] Liu B, Zhang N, Zhou C, *et al.* $(\text{La}_{0.25}\text{Sm}_{0.25}\text{Gd}_{0.25}\text{Ho}_{0.25})\text{AlO}_3$: High-entropy perovskite with harmonizing mechanical-thermal properties for TBCs. *J Am Ceram Soc* 2025, **108**: e70112.
- [14] Chen H, Zhao ZF, Xiang HM, *et al.* High entropy $(\text{Y}_{0.2}\text{Yb}_{0.2}\text{Lu}_{0.2}\text{Eu}_{0.2}\text{Er}_{0.2})_3\text{Al}_5\text{O}_{12}$: A novel high temperature stable thermal barrier material. *J Mater Sci Technol* 2020, **48**: 57–62.
- [15] Zhao ZF, Xiang HM, Chen H, *et al.* High-entropy $(\text{Nd}_{0.2}\text{Sm}_{0.2}\text{Eu}_{0.2}\text{Y}_{0.2}\text{Yb}_{0.2})_4\text{Al}_2\text{O}_9$ with good high temperature stability, low thermal conductivity, and anisotropic thermal expansivity. *J Adv Ceram* 2020, **9**: 595–605.
- [16] Majee BP, Bryce K, Huang LP, *et al.* A high-entropy rare-earth phosphate and its principle single component REPO_4 for environmental barrier coatings. *J Adv Ceram* 2025, **14**: 9221041.
- [17] Chang HQ, Liu PX, Lu H, *et al.* Discovery of intrinsically high-hardness high-entropy carbides. *Acta Mater* 2026, **309**: 122135.
- [18] Zhang KY, Du JL, Xiao JW, *et al.* Atomistic underpinnings for dislocation emission behaviors at the crack tips in FCC metals in light of thermo-kinetic synergy. *J Mater Inform* 2025, **5**: 49.
- [19] Almishal SSI, Furst M, Tan YZ, *et al.* Thermodynamics-inspired high-entropy oxide synthesis. *Nat Commun* 2025, **16**: 8211.
- [20] Zhang XM, Sun J, Zhang YY, *et al.* Machine learning for phase prediction of high entropy carbide ceramics from imbalanced data. *npj Comput Mater* 2026, **12**: 1–14.
- [21] Xu K, Shi XM, Xi ZC, *et al.* Machine-learning-guided design of incommensurate antiferroelectrics via field-driven phase engineering. *Adv Sci* 2026, **13**: e23873.
- [22] Yang HY, Gao C, Jiang DH, *et al.* Machine learning assisted prediction for the coefficient of thermal expansion of binary crystals. *J Adv Ceram* 2025, **14**: 9221117.
- [23] Pilania G, Mannodi-Kanakkithodi A, Uberuaga BP, *et al.* Machine learning bandgaps of double perovskites. *Sci Rep-UK* 2016, **6**: 19375.
- [24] Zhang J, Xu B, Xiong YX, *et al.* Design high-entropy carbide ceramics from machine learning. *npj Comput Mater* 2022, **8**: 5.
- [25] Fan Y, Bai YL, Li Q, *et al.* Compositional design and phase formation capability of high-entropy rare-earth disilicates from machine learning and decision fusion. *npj Comput Mater* 2024, **10**: 95.
- [26] Lin YL, Zou LX, Dai JW, *et al.* First-principles application in the research field of thermal barrier coating. *Adv Ceram* 2025, **46**: 24–51. (in Chinese)
- [27] Kresse G, Furthmüller J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys Rev B* 1996, **54**: 11169–11186.
- [28] Blöchl PE. Projector augmented-wave method. *Phys Rev B* 1994, **50**: 17953–17979.
- [29] Kresse G, Joubert D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys Rev B* 1999, **59**: 1758–1775.
- [30] Perdew JP, Ruzsinszky A, Csonka GI, *et al.* Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys Rev Lett* 2008, **100**: 136406.
- [31] Perdew JP, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Phys Rev Lett* 1996, **77**: 3865–3868.
- [32] Murnaghan FD. The compressibility of media under extreme pressures. *P Natl Acad Sci USA* 1944, **30**: 244–247.
- [33] Birch F. Finite elastic strain of cubic crystals. *Phys Rev* 1947, **71**: 809–824.
- [34] Wu B, Zinkevich M, Aldinger F, *et al.* *Ab initio* study on structure and phase transition of A- and B-type rare-earth sesquioxides Ln_2O_3 ($\text{Ln} = \text{La-Lu}$, Y, and Sc) based on density function theory. *J Solid State Chem* 2007, **180**: 3280–3287.
- [35] Ye BL, Wen TQ, Nguyen MC, *et al.* First-principles study, fabrication and characterization of $(\text{Zr}_{0.25}\text{Nb}_{0.25}\text{Ti}_{0.25}\text{V}_{0.25})\text{C}$ high-entropy ceramics. *Acta Mater* 2019, **170**: 15–23.
- [36] Pianassola M, Loveday M, Lalk R, *et al.* Crystal growth and phase formation of high-entropy rare-earth monoclinic aluminates. *J Am Ceram Soc* 2023, **106**: 7123–7132.
- [37] Chen XL, Chen L, Tian J, *et al.* Measurements and applications of thermal-mechanical properties of thermal/environmental barrier coating materials. *Adv Ceram* 2025, **46**: 389–415. (in Chinese)
- [38] Bao YW, Hu CF, Zhou YC. Damage tolerance of nanolayer grained ceramics and quantitative estimation. *Mater Sci Tech-Lond* 2006, **22**: 227–230.
- [39] Leitner J, Chuchvalec P, Sedmidubský D, *et al.* Estimation of heat capacities of solid mixed oxides. *Thermochim Acta* 2002, **395**: 27–46.
- [40] Chu KL, Zhang YN, Zhao JL, *et al.* Screening rare-earth aluminates as promising thermal barrier coatings by high-throughput first-principles calculations. *J Am Ceram Soc* 2023, **106**: 3089–3102.
- [41] Xia H, Sha ZD. Predicting thermophysical properties and fracture toughness of zirconia-based thermal barrier coating by interpretable machine learning method. *Mater Today Commun* 2025, **46**: 112842.
- [42] Lu JQ, Zhang FP, Wang WY, *et al.* Discovering superhard high-entropy diboride ceramics via a hybrid data-driven and knowledge-enabled model. *J Am Ceram Soc* 2023, **106**: 6923–6936.
- [43] Zhang Y, Wen C, Wang CX, *et al.* Phase prediction in high entropy alloys with a rational selection of materials descriptors and machine learning models. *Acta Mater* 2020, **185**: 528–539.
- [44] Huang WJ, Martin P, Zhuang HL. Machine-learning phase prediction of high-entropy alloys. *Acta Mater* 2019, **169**: 225–236.
- [45] Zhang J, Xiang XP, Xu B, *et al.* Rational design of high-entropy ceramics based on machine learning—A critical review. *Curr Opin Solid St M* 2023, **27**: 101057.
- [46] Zhou ZQ, Zhou YJ, He QF, *et al.* Machine learning guided appraisal and exploration of phase design for high entropy alloys. *npj Comput Mater* 2019, **5**: 128.
- [47] Breiman L. Random forests. *Mach Learn* 2001, **45**: 5–32.
- [48] Zhu XZ, Tsang DCW, Wang L, *et al.* Machine learning exploration of the critical factors for CO_2 adsorption capacity on porous carbon materials at different pressures. *J Clean Prod* 2020, **273**: 122915.
- [49] Marzban C. The ROC curve and the area under it as performance measures. *Weather Forecast* 2004, **19**: 1106–1114.

- [50] Antwarg L, Miller RM, Shapira B, *et al.* Explaining anomalies detected by autoencoders using Shapley Additive Explanations. *Expert Syst Appl* 2021, **186**: 115736.
- [51] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. California, USA 2017: 4768–4777.
- [52] Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 2019, **1**: 206–215.
- [53] Lundberg SM, Erion G, Chen H, *et al.* From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020, **2**: 56–67.
- [54] Zhang ZX, Hou CY, Zhang ZY, *et al.* Data-driven design of spinodal decomposition in (Ti,Zr,Hf)C composite carbides for optimizing the hardness-toughness trade-off. *Adv Funct Mater* 2025, **35**: 2502555.
- [55] Ma SL, Rao YC, Huang X, *et al.* High-throughput discovery of metal oxides with high thermoelectric performance *via* interpretable feature engineering on small data. *Mater Today Phys* 2024, **45**: 101457.
- [56] Clarke DR. Materials selection guidelines for low thermal conductivity thermal barrier coatings. *Surf Coat Tech* 2003, **163**: 67–74.
- [57] Chen L, Hu MY, Guo J, *et al.* Mechanical and thermal properties of RETaO₄ (RE = Yb, Lu, Sc) ceramics with monoclinic-prime phase. *J Mater Sci Technol* 2020, **52**: 20–28.