

Research Article

Open Access



# Prediction of temperature-dependent yield strength of refractory high entropy alloy based on stacking integrated framework

Liping Yu<sup>1</sup>, Jiarui Zhai<sup>2</sup>, Wenzhe Cao<sup>2</sup>, Jingli Ren<sup>2,\*</sup>

<sup>1</sup>School of Artificial Intelligence and Big Data, Henan University of Technology, Zhengzhou 450001, Henan, China.

<sup>2</sup>School of Mathematics and Statistics, Zhengzhou University, Zhengzhou 450001, Henan, China.

**\*Correspondence to:** Prof. Jingli Ren, School of Mathematics and Statistics, Zhengzhou University, No.100, Science Avenue, Hi-Tech Zone, Zhengzhou 450001, Henan, China. E-mail: renjl@zzu.edu.cn

**How to cite this article:** Yu L, Zhai J, Cao W, Ren J. Prediction of temperature-dependent yield strength of refractory high entropy alloy based on stacking integrated framework. *J Mater Inf* 2024;4:28. <http://dx.doi.org/10.20517/jmi.2024.39>

**Received:** 26 Aug 2024 **First Decision:** 31 Oct 2024 **Revised:** 21 Nov 2024 **Accepted:** 28 Nov 2024 **Published:** 17 Dec 2024

**Academic Editor:** Chuang Dong **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

## Abstract

Refractory high-entropy alloys (RHEAs) are promising materials for high-temperature applications. This study introduces an interpretable prediction model for the temperature-dependent yield strength of these alloys, utilizing a stacking ensemble algorithm. The Kolmogorov-Smirnov (K-S) test results ( $P = 0.873$ ) confirm the model's reliability. Key features are analyzed from global and local perspectives using accumulated local effect (ALE) and SHapley Additive exPlanations (SHAP) methods. The analysis reveals that an atomic size difference ( $\delta > 0.049$ ) and a bulk modulus ( $K > 150$ ) positively influence yield strength. Bivariate partial dependence plots (PDP) demonstrate that yield strength increases with both  $K$  and the shear modulus ( $G$ ). At high temperatures (1,000 and 1,200 °C), the stacking model, combined with the Dung Beetle optimization algorithm, predicts improved yield strength in various alloy compositions. For example, the alloy  $\text{Cr}_{0.275}\text{Nb}_{0.215}\text{Mo}_{0.349}\text{V}_{0.161}$  shows a 19.90% improvement compared to the original dataset. Using parallel coordinate plots and iterative analysis, promising concentration regions for high yield strength were identified, such as in the Al-Cr-Nb-Mo-V system, where lower Al and Nb content and higher Cr content enhance performance.

**Keywords:** Refractory high-entropy alloy, temperature-dependent yield strength prediction, stacking ensemble algorithm, model interpretation, dung beetle optimization algorithm



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



## INTRODUCTION

High-entropy alloys (HEAs) are a unique class of materials composed of four or more elements in equimolar or near-equimolar ratios<sup>[1,2]</sup>. Their multi-element composition imparts exceptional properties, including high strength, hardness, corrosion resistance, and oxidation resistance<sup>[3–6]</sup>. These attributes have made HEAs a significant focus of materials research, especially for applications requiring superior mechanical performance. Building on the foundation of HEAs, refractory HEAs (RHEAs) were introduced in 2010 to further enhance high-temperature performance<sup>[7]</sup>. Early studies demonstrated the promise of RHEAs, such as NbMoTaWV and NbMoTaW, which exhibited hardness values 2–3 times higher than predicted by theoretical calculations<sup>[8]</sup>. Over time, the compositional space of RHEAs expanded to include lighter elements such as Al, further enhancing their versatility and applications.

The demand for materials capable of maintaining performance under extreme conditions has increased significantly with advancements in aerospace, power generation, and petrochemical industries. While nickel-based superalloys have been the traditional choice for high-temperature applications, they face challenges such as insufficient mechanical properties, susceptibility to corrosion, and cracking<sup>[9]</sup>. RHEAs have emerged as a superior alternative due to their exceptional thermal stability and mechanical strength. For example, Steingrimsson *et al.* demonstrated that certain RHEAs outperform nickel-based superalloys in strength using a bilinear logarithmic model<sup>[10]</sup>. Similarly, Senkov *et al.* developed  $\text{AlMo}_{0.5}\text{NbTa}_{0.5}\text{TiZr}_x$  ( $x = 0.5, 1$ ) with yield strengths of 1,870 and 1,597 MPa at 600 and 800 °C, respectively<sup>[2,11,12]</sup>.  $\text{AlMo}_{0.5}\text{NbTa}_{0.5}\text{TiZr}_{0.5}$  achieved an impressive 935 MPa at 1,000 °C nearly four times the yield strength of conventional nickel-based superalloys. These advancements highlight the potential of RHEAs for high-temperature applications where stability and strength are critical.

However, the vast compositional space of RHEAs presents both opportunities and challenges. While this flexibility enables tailored property optimization, it also complicates alloy design. Traditional approaches, such as molecular dynamics simulations, density functional theory, and phase diagram calculations, are costly and time-consuming. The rise of machine learning (ML) offers a promising alternative for alloy property prediction by enabling data-driven insights and rapid evaluations<sup>[13–19]</sup>. For instance, Rickman *et al.* used a genetic algorithm and multiple linear regression (LR) to discover high-hardness alloys, while Xiong *et al.* developed a random forest (RF) model to estimate Vickers hardness and tensile strength<sup>[20,21]</sup>. Additionally, Khatavkar *et al.* employed Gaussian process regression (GPR) to predict creep rupture life and assess feature significance in superalloys<sup>[22]</sup>. Although ML has been applied to address several challenges in materials science, there has been limited research on predicting temperature-dependent yield strength, specifically for RHEAs.

In this study, we develop a ML model tailored for the prediction of temperature-dependent yield strength in RHEAs. Unlike prior studies, our approach emphasizes predictive stability and accuracy through ensemble learning (EL) techniques. We leverage prior knowledge to identify key thermodynamic parameters that influence yield strength and use both local and global interpretability methods, such as accumulated local effects (ALE) and SHapley additive exPlanations (SHAP), to analyze the impact of individual and paired features. Furthermore, we employ the Dung Beetle optimization algorithm to identify optimal alloy compositions with enhanced high-temperature performance. This integrative approach not only improves predictive accuracy but also provides actionable insights for designing high-performance RHEAs.

## MATERIALS AND METHODS

### Database establishment

This study focuses on RHEA systems, which consist of nine high melting point elements (Zr, Ti, Hf, Nb, Mo, V, Ta, W and Cr) and one light element (Al). Data on the temperature-dependent yield strength and compositions of these alloys were compiled from multiple published studies<sup>[7,23–26]</sup>. The data were selected

to encompass a wide range of chemical compositions by including alloys with varying elemental compositions and ratios. To capture the microstructural diversity of RHEAs, alloys with different grain sizes, morphologies, distributions, and phase compositions were also considered. The data were preprocessed to eliminate duplicate values, outliers, and entries that were broken within the elastic strain range. After preprocessing, the dataset consisted of 275 data points: 54 ternary alloys, 168 quaternary alloys, and 53 quinary alloys. Given the empirical criterion for testing data selection, the dataset is separated into sets for training and testing in a 4:1 ratio.

The selected feature is closely linked to the target attribute, ensuring that the model's predicted values are close to the real values within an acceptable error range. According to Ref. [21], the mechanical properties of high entropy alloys are highly dependent on their phases, such as hardness, yield strength, ultimate tensile strength, and so on. Hou *et al.* studied the relationship between the solid solution phase and valence electron concentration (VEC), mixing entropy ( $\Delta S_{\text{mix}}$ ), mixing enthalpy ( $\Delta H_{\text{mix}}$ ), atomic radius difference ( $\delta$ ), electronegative difference ( $\Delta\chi$ ),  $\Omega$  and other parameters of high entropy alloy, and established a model to solve the phase classification problem [27]. The bulk modulus (K) and shear modulus (G) are directly connected to the alloy's mechanical characteristics [28]. K calculates the fracture resistance and demonstrates the degree to which the alloy's volume changes in response to external forces. G measures the extent of shear deformation and evaluates the propensity of fracture resistance to rise following the commencement of plastic deformation. Moreover, increasing the experimental temperature enhances atomic diffusion capacity, increases the number of vacancies in the material, and alters the grain boundary slip system, leading to significant fluctuations in strength. As a result, it is considered an important feature [29].

Supplementary Table 1 summarizes the 21 descriptors developed in the current study. The properties of each element were taken from the appropriate sources [2], and the enthalpy of mixing between the two elements was determined from a table in the literature [30]. Because there is no evident distinction between “solvent” and “solute” in HEAs, the average weighting of elemental characteristics is used to define this random solid solution, as given in

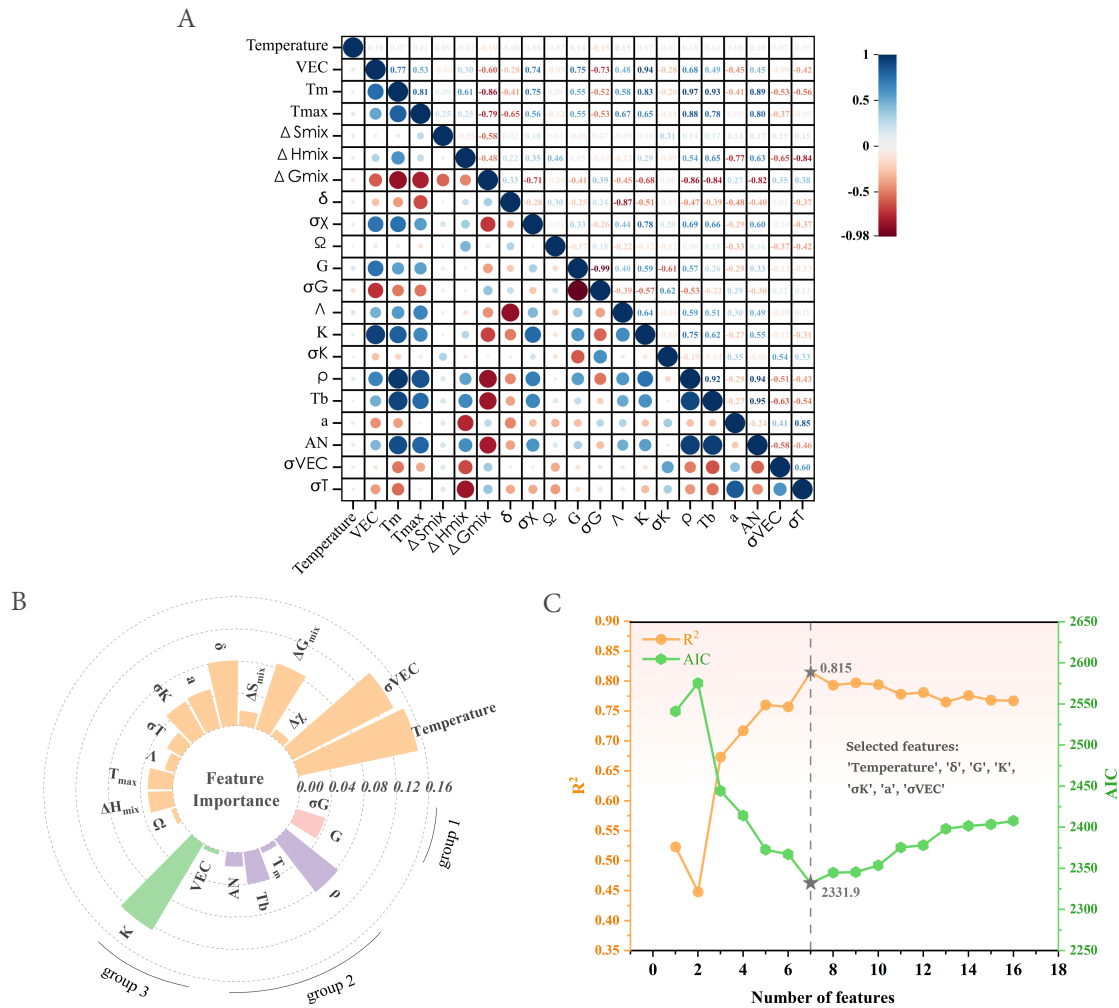
$$P = \sum_{i=1}^n c_i P_i \quad (1)$$

where the alloying element proportion is expressed by  $c_i$  and the related properties by  $P_i$ . Considering the tendency of some features, such as VEC, to exhibit strong localization, the “concentration mean difference” is calculated here to represent the combined effect of elemental property mismatching, as given in

$$\sigma P = \sqrt{c_i (P_i - \bar{P})^2} \quad (2)$$

where  $\bar{P}$  is the mean value.

In the process of model training, it is necessary to convert the original data into dimensionless index evaluation values. This is because large numbers can cause numerical issues, and when the model uses gradient descent, elliptical contours result in more iterations. Therefore, some algorithms require scaling of eigenvalues, such as support vector machine, integrated learning, K-nearest neighbors (KNN), *etc.* We use Z-score feature



**Figure 1.** Feature selection steps: (A) heat map of Pearson's correlation coefficient; (B) lollipop plot of feature importance; (C) RFE process with AIC and R as evaluation indexes. RFE: Recursive feature elimination; AIC: Akaike information content.

normalization to standardize the data:

$$x' = \frac{x - \mu}{\sigma} \quad (3)$$

where  $x$  is the original value of the data,  $x'$  is the normalized feature value,  $\mu$  is the mean of the population data, and  $\sigma$  is the standard deviation of the population data.

### Feature screening

Although an excessive number of distinctive features offers a physical foundation for yield strength prediction, feature duplication reduces physical insight into the problem and raises the likelihood of overfitting. As a result, feature screening is required to identify the most representative feature subset, thereby improving the model's interpretability to some extent [31,32].

#### Pearson correlation coefficient and feature importance

The presence of a large number of highly correlated characteristics causes multicollinearity, which means that they are assigned greater weights, passively decreasing the information contained in other features and so compromising the regression model's prediction accuracy. The Pearson correlation coefficient (PCC) value measures the degree of correlation between any two features and ranges from -1 to 1. The closer the PCC

value is to 1, the higher the similarity between the features. The PCC value is calculated by:

$$\text{PCC} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4)$$

where the numerator is the covariance of features, and the denominator is the product of the standard deviations of the two features. If the absolute value of the correlation coefficient is greater than 0.9, it implies that the two characteristics have a strong association<sup>[33]</sup>. Figure 1A illustrates the PCC heat map, highlighting three sets of correlated features: [(“VEC”, “K”), (“T<sub>m</sub>”, “ρ”, “Tb”, “AN”) and (“G”, “σG”)].

By calculating the PCC values, we have taken into account the relationship between the variables. The next step is to fit a predictive model using the dataset, establish the logical relationship between features and the target attribute, and measure the relative contribution of each input feature using feature importance. This process helps effectively remove variables with low values. Figure 1B visualizes the importance of each feature, with values decreasing from left to right, where perpendiculars of the same color represent combinations of features with high correlation. Features with low importance (“VEC”, “T<sub>m</sub>”, “σG”, “Tb”, “AN”) were removed. Finally, 16 relatively independent features with great influence on the target were selected from the original 21 features by PCC value and feature importance.

#### *Recursive feature elimination*

This study employed recursive feature elimination (RFE) combined with cross-validation (CV) to further reduce the dimensionality of variables and enhance the model's computational efficiency. The model was constructed multiple times, with weak features being eliminated based on performance evaluation indices, until only one feature remained.

The evaluation indexes selected here are respectively  $R^2$  and Akaike information content (AIC) criteria<sup>[34,35]</sup>.  $R^2$  evaluates the degree of the model fitting to the sample, and AIC introduces a penalty term on this basis to remove features as much as possible, that is, select a simple model with few parameters to prevent the occurrence of overfitting. The calculation is written as follows:

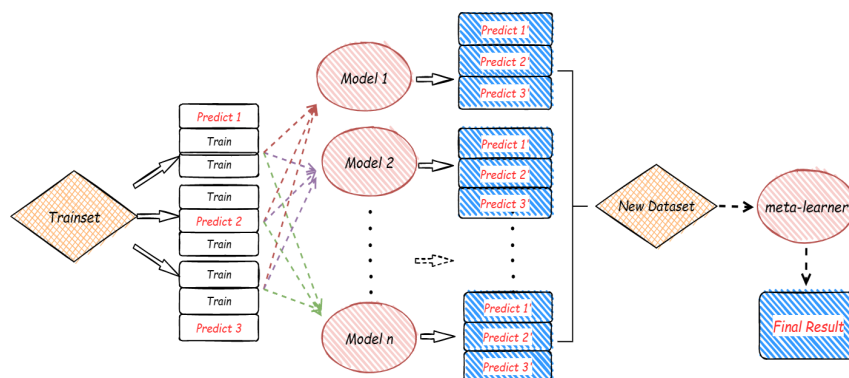
$$\text{AIC} = 2k - 2 \ln(L) = 2k + n \ln(\text{MSE}) \quad (5)$$

where  $k$  is the number of variables,  $L$  is the likelihood function, and MSE is the mean-square error. A lower  $k$  indicates a simpler model, whereas a higher  $L$  indicates a more accurate model. In general, the lower the AIC, the better the model's overall performance. As shown in Figure 1C, when the remaining seven characteristics are “Temperature”, “δ”, “G”, “K”, “σK”, “a” and “σVEC”,  $R^2$  reaches a maximum of 0.815 and AIC reaches a minimum of 2,331.9.

#### **Model construction**

To select a model with strong fit and generalization performance, we split the original dataset into a training set and a testing set in a 4:1 ratio, using key features as input. We then applied several classic ML algorithms including extreme gradient boosting regressor (XGBR), extremely randomized trees (ERT), RF, adaptive boosting (AdaBoost), GPR, KNN, and LR to the training set for repeated model building. Preliminary modeling, summarized in Supplementary Table 2, indicated that the XGBR model provided superior predictive accuracy, suggesting that gradient boosting algorithms are well-suited for predicting temperature-dependent yield strength. Consequently, we included additional models in the gradient boosting framework, such as histogram-based gradient boosting regressor (HistGBR) and LightGBM regressor (LGBMR).

While traditional ML models effectively handle structured data, their noise resistance and extrapolation capabilities are often limited. The stacking EL algorithm, based on the meta-learner concept, addresses this by



**Figure 2.** The flow of the stacking algorithm.

combining multiple ML algorithms over two or more stages to produce a model with lower classification or regression errors. This progressive, optimal learning approach has been effective in various applications, such as predicting glass formation in amorphous metals<sup>[36]</sup>, detecting sensor faults<sup>[37]</sup>, and supporting hydropower emergency responses<sup>[38]</sup>. Therefore, we leverage the EL model to improve predictive accuracy in this study.

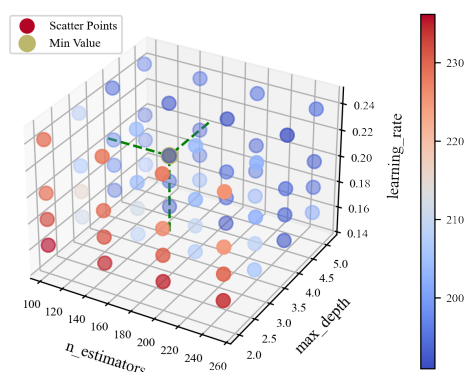
HEAs have complex compositions, and the interactions among the various elements can significantly influence the alloy's properties. One key advantage of the stacked integration framework is its ability to leverage the strengths of multiple meta-learners, enabling it to capture diverse features in the data. This approach allows for a more comprehensive consideration of factors such as alloy composition, structure, temperature, and other variables that affect yield strength, thereby enhancing prediction accuracy. In contrast, traditional methods often rely on empirical formulas derived from limited experimental data, which may not be applicable across the full range of alloy compositions and temperature conditions. Additionally, physical models may be overly simplified and fail to accurately capture the intricate interactions and microstructural changes within the alloys. Furthermore, studies<sup>[39–41]</sup> also demonstrated that the stacked integration framework yielded significant improvements in predicting the properties of RHEAs, offering a more reliable tool for the design and optimization of these materials.

The EL procedure involves using the highest-performing model as the base learner, training each independently, and then stacking their outputs to form a new dataset with dimensions  $(a, b)$ , where  $a$  is the sample count and  $b$  the number of base learners. The meta-learner then fits this transformed dataset to produce the final prediction. To prevent overfitting and manage multicollinearity, we selected simple models for the meta-learner, such as Lasso or linear regression, with regularization techniques, which have proven effective across fields, including sparse robust estimation<sup>[42]</sup>, multi-core learning prediction<sup>[43]</sup>, and feature selection and clustering<sup>[44]</sup>. Figure 2 shows the flow diagram of the integrated learning algorithm.

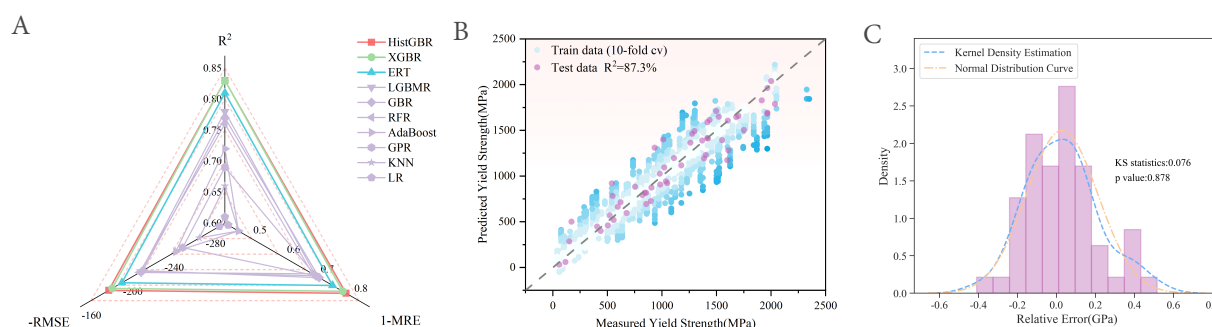
The use of stacking EL algorithm, which integrates multiple models, can increase prediction complexity and reduce transparency. In practice, when the output deviates from expected results, it can be difficult to isolate which specific model or input variable is responsible, especially when considering the complex interactions between various alloying elements and phases in RHEAs. This lack of interpretability can undermine the model's usefulness in explaining specific material behaviors. For instance, a prediction of yield strength under extreme temperatures might be affected by phase instability or grain boundary weakening, which may not be adequately captured by the integrated model.

In ML, overfitting where the model performs exceptionally well on training data but poorly on unseen data can hinder generalization. To mitigate this, we use  $k$ -fold CV to evaluate the model's generalization capacity, tuning parameters via grid search. Here,  $k$  is set to 10, as this value minimizes bias in error estimation<sup>[45]</sup>.





**Figure 3.** Grid search results for three hyperparameters of XGBR model. XGBR: Extreme gradient boosting regressor.



**Figure 4.** Model validation: (A) radar plot for CV; (B) scatter plot of model fit on the training and testing sets; (C) test set relative error distribution diagram. CV: Cross-validation.

Repeating 10-fold CV ten times further stabilizes our findings and improves reliability. We selected the coefficient of determination ( $R^2$ ), root mean square error (RMSE), and mean relative error (MRE) as evaluation metrics. MRE, an L1-norm loss, encourages feature sparsity, while RMSE, an L2-norm loss sensitive to outliers, promotes feature selection to enhance generalization. Together, MRE and RMSE complement each other in balancing model performance.

## RESULTS AND DISCUSSION

### Model comparison

Hyperparameter optimization plays a critical role in improving model performance. In this study, grid search was employed to fine-tune the hyperparameters of each ML model. For instance, with the XGBR model, hyperparameters A, B, and C were explored across a grid of 2,640 possible combinations. Figure 3 illustrates the final XGBR model in three dimensions, displaying max depth, learning rate, and  $n$  estimators. Supplementary Table 3 lists the optimized hyperparameters for each model.

Model performance was evaluated using 10-fold CV on the training set, with metrics including the  $R^2$ , RMSE, and MRE. Figure 4A shows the results, where RMSE is presented in its negative form, and MRE is expressed as 1-MRE for consistency. Coverage areas represent each method, where a larger area indicates better model performance. HistGBR and XGBR exhibited superior coverage areas, encapsulating the performance of other models. ERT stood out for its strong RMSE values. Consequently, HistGBR, XGBR, and ERT were selected for further investigation.

Following the classical ML model analysis, we examined model performance using EL. A stacking model was

constructed by pairing the three selected methods and applying the lasso algorithm as a meta-learner for ten times 10-fold CV on the training data [Supplementary Figure 1]. Compared to individual ML models, EL models showed higher  $R^2$  and 1-MRE values, lower RMSE values, and shorter error bars, indicating superior accuracy and reliability. Specifically, the ERT and HistGBR combination produced a stacking model with smaller error bars for  $R^2$ , 1-MRE, and RMSE, demonstrating more consistent performance and making it the final model selection. The model's predictive performance was then evaluated on the test set [Figure 4B]. Most data points are closely aligned with the diagonal, and the  $R^2$  on the test set is 0.873, demonstrating strong agreement between predicted and experimental values.  $R^2$  value of 0.87 indicates that 87% of the variance in the material properties is explained by the model, suggesting a strong correlation between predicted and observed data. Statistically, this suggests that 87% of the variance in the dependent variable (e.g., material properties such as yield strength, hardness, etc.) can be explained by the model's independent variables.

Higher  $R^2$  values, in practice, can help mitigate issues such as material wastage, production delays, and safety incidents caused by inaccurate predictions. They can also be leveraged to optimize designs and processes, enhancing performance and reducing costs. A high-precision predictive model enables engineers to more accurately assess and optimize performance, effectively screen and refine the composition and heat treatment processes of RHEAs, and ultimately advance the use of these alloys in high-temperature and more demanding environments.

### Model evaluation

#### *Residual normality test*

To evaluate the reliability and accuracy of the ERT-HistGBR-based model, residual analysis was conducted on the training set. Regression models assume that residuals follow a normal distribution, maintaining maximum uncertainty in the error residual. If this assumption holds, it implies random errors without systematic bias, making the model suitable for prediction.

Figure 4C presents the results of the residual normality test using a histogram and the Kolmogorov-Smirnov (K-S) test. The histogram shows a close alignment between the kernel density fit (blue dashed line) and the normal distribution curve (red dashed line). The K-S test yielded a  $P$ -value of 0.076 ( $> 0.05$ ), supporting the null hypothesis that the residuals follow a normal distribution at a 95% confidence level.

#### *Reliability and applicability test*

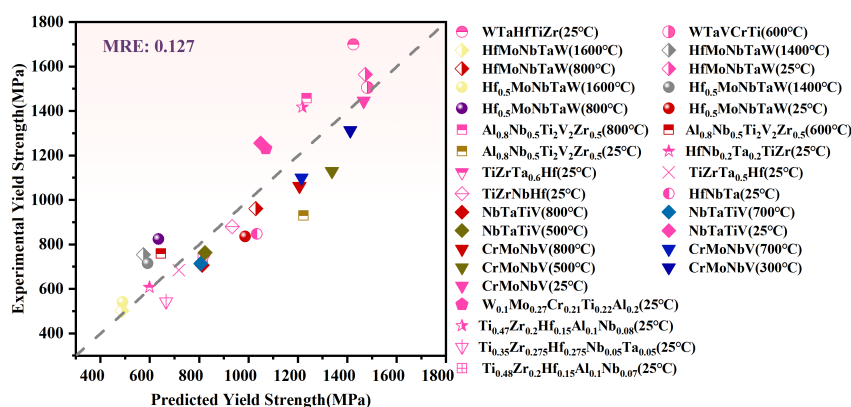
The robustness of the stacking model was further evaluated using 31 external data points from diverse sources [46–50]. Figure 5 illustrates the deviation between predicted and experimental values, where different symbols represent alloys, and matching colors indicate similar temperature conditions. Data points cluster near the diagonal, indicating strong predictive performance across varying alloy compositions and temperature conditions. To benchmark the stacking model, several classical ML algorithms, including XGBR, RFR, AdaBoost, and LR, were applied to the same external dataset. The stacking model consistently outperformed these models in terms of MRE, as summarized in Supplementary Figure 2.

Feature generalizability was also tested using 59 temperature-dependent yield strength data points from HEAs containing Si, Co, Fe, Ni, and Mn. Using features such as “Temperature”, “ $\delta$ ”, “G”, “K”, “ $\sigma_K$ ”, “a”, “ $\sigma_{VEC}$ ”, we built a stacking model for prediction, averaging results from 5-fold CV. Supplementary Figure 3 provides model predictions, achieving a  $R^2$  of 0.824. This suggests that our selected features generalize well in predicting yield strength for other RHEA systems.

### Model interpretability analysis

ML models, often viewed as black boxes, can obscure the mechanisms behind their predictions, limiting insights into the mechanical characteristics of RHEAs. To address this, our study uses interpretability strategies





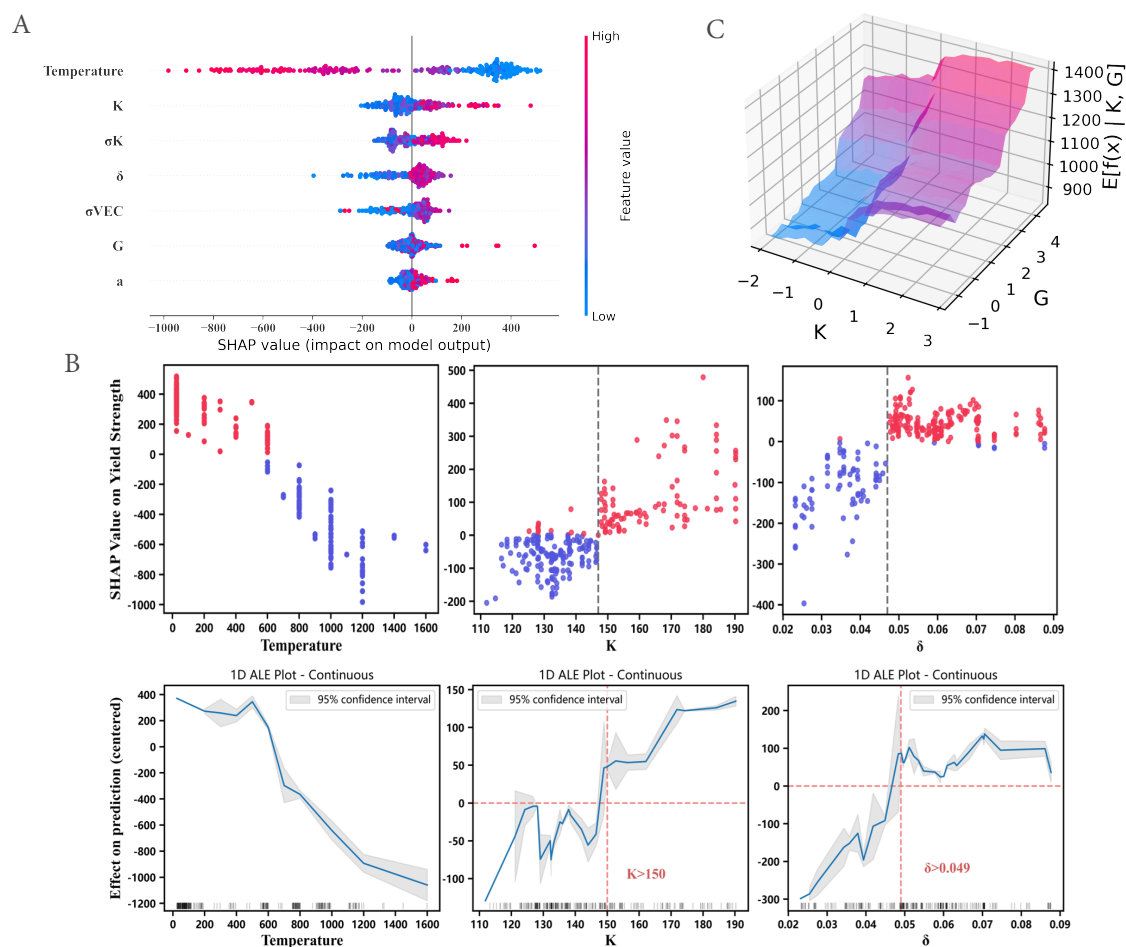
**Figure 5.** Fitted scatterplot for experimental data by the model based on ERT and HistGBR. ERT: Extremely randomized trees; HistGBR: histogram-based gradient boosting regressor.

to analyze the effects of various attributes on prediction outcomes, categorizing methods as global or local based on their focus. Global methods, including partial dependence plots (PDP) [51] and ALE [52], show the average influence of features across the data distribution. Local methods, such as SHAP [53–55] and individual conditional expectation (ICE) [52], focus on individual instances.

Here, we used the pdpbox package to generate PDP and ICE plots, with PDP showing the average dependency between yield strength and selected features, while ICE plots illustrate individual sample effects, minimizing confounding factors. [Supplementary Figure 4](#) shows that all of the ICE curves are consolidated at a single starting point, making it easier to understand the connection between local and global effects. By analyzing the features from both perspectives, it becomes clear that temperature has a negative impact on yield strength, and the influence of each feature on individual samples generally follows the overall average trend. As temperature increases, the material's lattice structure changes, and atomic thermal vibrations intensify, leading to a negative correlation between temperature and the target properties. This, in turn, enhances the material's plastic deformation ability while reducing its yield strength.

To further investigate, SHAP values and ALE plots are employed, with the average effect line in the ALE plots expected to lie within the gray confidence interval at a 95% confidence level. Figure 6A shows the SHAP summary plots, which rank the importance of the model's seven attributes. The temperature distribution across the samples is highly spread, indicating that this feature has the most significant influence on the prediction. The interaction between the lattice constant  $a$  and other features has no notable effect, which is attributed to its low importance. Therefore, taking temperature, atomic size difference  $\delta$  and bulk modulus  $K$  as research parameters, Figure 6B illustrates that both from a single point perspective and an average effect standpoint, the atomic size difference  $\delta$  and bulk modulus  $K$  exhibit a critical point, marking a clear boundary between two distinct groups in the plots.

By analyzing the intersection of the interval defined by this boundary, we find that  $\delta > 0.049$  positively influences yield strength, while  $\delta < 0.049$  has a negative effect. The conclusion aligns with the findings of Wang *et al.*, which shows when  $\delta$  increases, HEAs are more likely to form a high-strength BCC/B2 phase<sup>[36]</sup>. BCC-type HEAs generally have a high yield strength. Similarly, the SHAP plot for the volume modulus K indicates that when  $K > 150$  GPa, it contributes positively to yield strength, and vice versa, which supports the findings of Lee *et al.* that an increase in the alloy's K value is often associated with improved fracture resistance during plastic deformation<sup>[56]</sup>. Furthermore, when a univariate effect is not significant, we explore the impact of bivariate interactions on the target property using a PDP three-dimensional diagram. As shown in Figure 6C, high values of both bulk modulus K and shear modulus G have a beneficial effect on yield strength. This finding is consis-



**Figure 6.** Model validation: (A) SHAP value distributions for key features of the stacking model. The data points represent individual alloy samples, with each point color-coded according to the magnitude of the relevant feature. Positive SHAP values indicate that the feature increases the yield strength, while negative SHAP values suggest a decrease in yield strength due to the feature; (B) SHAP dependence scatter plot and ALE plot for single-feature  $\delta$  and  $K$ ; (C) 3D partial dependence graph of interaction feature  $[K, G]$ . SHAP: SHapley Additive exPlanations; ALE: accumulated local effect.

tent with previous studies<sup>[57]</sup> that a combination of high bulk and shear modulus increases the yield strength for MoNbTaTi refractory complex concentrated alloys. The combination of a high bulk modulus and high shear modulus can enhance the yield strength of RHEAs through several mechanisms, including improved cohesion, increased resistance to dislocation slip, enhanced deformation resistance, and greater stability of the alloy's crystal structure.

### Yield strength optimization

Optimizing RHEA compositions for high yield strength at elevated temperatures is critical for material development. This study combined the stacking model with the dung beetle optimizer (DBO), a recent swarm intelligence algorithm by Xue *et al.*<sup>[58]</sup>. DBO thoroughly explores the solution space to avoid local optima, showing robust performance. As confirmed by Giles *et al.*, yield strength is temperature-dependent, making independent searches for optimal compositions at 1,000 and 1,200 °C effective<sup>[59]</sup>.

Table 1 presents optimized alloy compositions. For CrNbMoV and AlCrNbMoV systems at 1,000 °C, which exhibit baseline strengths of 1,036 and 1,085 MPa, respectively, optimization yielded  $\text{Al}_{0.050}\text{Cr}_{0.330}\text{Nb}_{0.060}\text{Mo}_{0.337}\text{V}_{0.223}$  with a strength of 1,300.638 MPa, representing a 264.638 MPa increase, primarily due to a rise in Mo content.

**Table 1. The results of composition optimization by DBO**

Composition	Temperature	Yield strength (MPa)	Improvement ratio
Al <sub>0.050</sub> Cr <sub>0.330</sub> Nb <sub>0.060</sub> Mo <sub>0.337</sub> V <sub>0.223</sub>	1,000	1,300.638	19.87%
Cr <sub>0.275</sub> Nb <sub>0.215</sub> Mo <sub>0.349</sub> V <sub>0.161</sub>	1,000	1,242.259	19.90%
Cr <sub>0.181</sub> Ta <sub>0.190</sub> Ti <sub>0.050</sub> W <sub>0.301</sub> V <sub>0.278</sub>	1,200	886.224	18.16%
Cr <sub>0.256</sub> Ta <sub>0.238</sub> W <sub>0.240</sub> V <sub>0.266</sub>	1,200	1,016.648	3.84%

DBO: Dung beetle optimizer.

Similarly, the optimized composition Cr<sub>0.275</sub>Nb<sub>0.215</sub>Mo<sub>0.349</sub>V<sub>0.161</sub> achieved 1,242.259 MPa, a 206.259 MPa again. At 1,200 °C, optimized alloys within Cr-Ta-Ti-W-V, Cr-Ta-W-V systems yielded Cr<sub>0.181</sub>Ta<sub>0.190</sub>Ti<sub>0.050</sub>W<sub>0.301</sub>V<sub>0.278</sub> and Cr<sub>0.256</sub>Ta<sub>0.238</sub>W<sub>0.240</sub>V<sub>0.266</sub>, showing yield strength improvements of 18.16% and 3.84%, compared to the original data.

### Effect of element content on yield strength

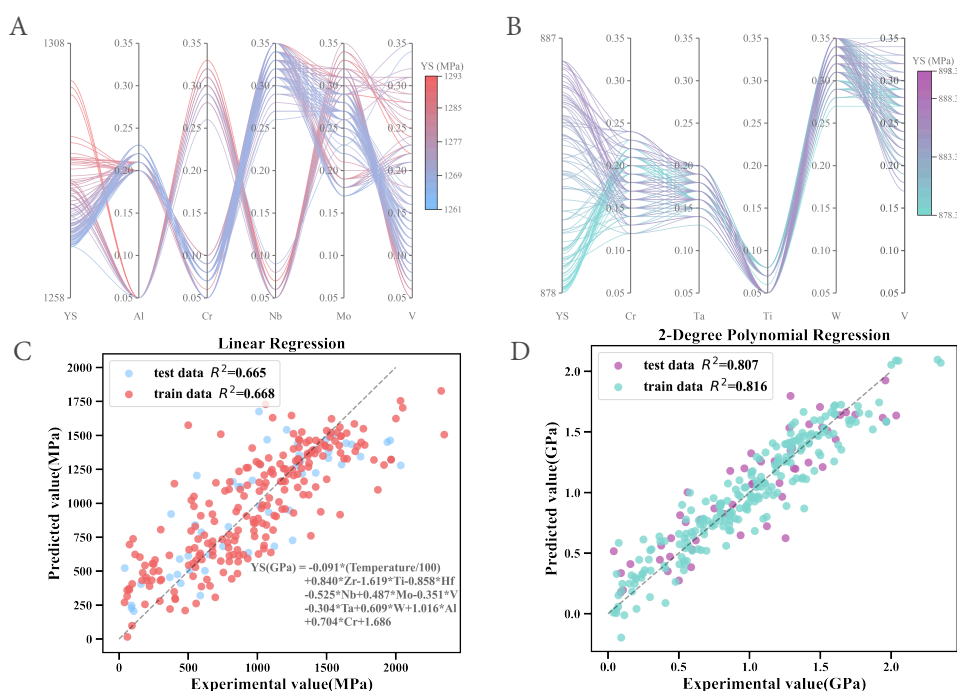
After multiple DBO cycles scanning the composition space, a dataset was generated, detailing the compositions and corresponding yield strengths of various RHEAs. This dataset facilitates analysis of the relationship between element ratios and performance. Given the complexity of RHEA systems, each alloy system was analyzed individually, with results for different temperatures and systems combined into four datasets. To reduce model error impact, we sorted the data by yield strength, selecting the top 200 entries exceeding the original alloy's target. Alloy compositions were rounded to two decimal places to avoid duplication.

Figure 7A illustrates the relationship between elements and yield strength for the Al-Cr-Nb-Mo-V alloy at 1,000 °C. Two distinct categories emerge: Mo concentration ranges from 17% to 35% for high yield strength, with Al(at%) = 5%, Cr(at%) = 26%-33%, and Nb(at%) = 5%-9%. V shows a broader distribution without a specific range. Thus, high Cr and low Al and Nb concentrations favor stronger alloys in the Al-Cr-Nb-Mo-V system at 1,000 °C. Similarly, Figure 7B shows the CrTaTiWV system at 1,200 °C, where high performance is achieved with Cr(at%) = 15%-23%, Ta(at%) = 16%-21%, Ti(at%) = 5%, W(at%) = 26%-35%, and V(at%) = 17%-35%. Supplementary Figure 5 provides parallel coordinates for the CrNbMoV system at 1,000 °C and the CrTaWV system at 1,200 °C, revealing similar ranges that favor yield strength.

Additionally, we examined the overall impact of nine RHEA elements. Using a dataset of 276 experimental samples, linear and quadratic polynomial regressions were applied, with temperature included as a feature. The dataset was split 4:1 into training and testing sets. Figure 7C displays the linear regression model, which achieved an R<sup>2</sup> score of 0.665 on the test set, indicating a clear relationship between yield strength and elemental content and temperature. This linear model offers a basic understanding of each element's effect. In contrast, the quadratic polynomial model achieved a better R<sup>2</sup> score of 0.807 [Figure 7D], highlighting its superior predictive ability and suitability for quantitative yield strength predictions. Coefficients and intercepts for the quadratic terms are detailed in Supplementary Table 4.

## CONCLUSION

This work integrates correlation analysis, feature ranking, and RFE to select seven important features, including temperature,  $\delta$ , G, K,  $\sigma_K$ ,  $a$  and  $\sigma_{VEC}$ . To enhance the accuracy of temperature-dependent yield strength predictions for RHEAs, HistGBR and ERT models are combined through a stacking method. The developed stacking model is validated as plausible using the residual analysis technique (K-S test), and the model's R<sup>2</sup> value on the independent test set is 0.873. Furthermore, the combination of the local interpretability approach ALE and the global interpretability method SHAP demonstrates that the univariate atomic size difference and bulk modulus have clear critical values; i.e., when  $\delta > 0.049$  or  $K > 150$ , they positively contribute to the goal attributes. The PDP three-dimensional figure demonstrates that the yield strength increases when the



**Figure 7.** Parallel coordinates plots show the values of the elements of (A) AlCrNbMoV and (B) CrTaTiWV system; Fitted scatter plot of (C) linear regression; (D) quadratic polynomial regression.

bivariable K and shear modulus G increase.

Using the model of coupling intelligent optimization algorithm DBO iterative search at 1,000 and 1,200 °C, we extract better yield strengths of the alloy composition proportion. Compared to the result of the original dataset, alloys  $\text{Al}_{0.050}\text{Cr}_{0.330}\text{Nb}_{0.060}\text{Mo}_{0.337}\text{V}_{0.223}$  and  $\text{Cr}_{0.275}\text{Nb}_{0.215}\text{Mo}_{0.349}\text{V}_{0.161}$  have a yield strength increase rate of 19.87% and 19.90% based on computational predictions, respectively. Finally, applying the parallel coordinate plot (PCP) approach in conjunction with the supplemental data acquired by the DBO iteration, the linear and nonlinear effects of various system factors on yield strength at high temperatures are obtained.

## DECLARATIONS

### Authors' contributions

Investigation, visualization, writing - review and editing: Yu L

Data curation, visualization, writing - original draft: Zhai J

Visualization, writing editing: Cao W

Methodology, visualization, supervision, writing - review and editing, funding acquisition: Ren J

### Availability of data and materials

The detailed materials and methods used in the experiment are provided in the [Supplementary Materials](#). Other raw data that support the findings are available from the corresponding author upon reasonable request.

### Financial support and sponsorship

The authors would like to acknowledge the financial support from the NSFC (Nos. U23A2065, 52071298), the Natural Science Foundation of Henan Province (No. 232300420346), Key Research Programs of Higher Education Institutions in Henan Province (25A110001), and the Training Plan for Young Backbone Teachers of Henan University of Technology.

### Conflicts of interest

Ren J is the guest editor of the Special Issue, while the other authors have declared that they have no conflicts of interest.

### Ethical approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Copyright

© The Author(s) 2024.

## REFERENCES

1. Yeh JW, Chen SK, Lin SJ, et al. Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes. *Adv Eng Mater* 2004;6:299–303. DOI
2. Miracle DB, Senkov ON. A critical review of high entropy alloys and related concepts. *Acta Mater* 2017;122:448–511. DOI
3. Li W, Xie D, Li D, Zhang Y, Gao Y, Liaw PK. Mechanical behavior of high-entropy alloys. *Prog Mater Sci* 2021;118:100777. DOI
4. Nair RB, Arora HS, Mukherjee S, Singh S, Singh H, Grewal HS. Exceptionally high cavitation erosion and corrosion resistance of a high entropy alloy. *Ultrason Sonochem* 2018;41:252–60. DOI
5. Gorr B, Azim M, Christ HJ, Mueller T, Schliephake D, Heilmaier M. Phase equilibria, microstructure, and high temperature oxidation resistance of novel refractory high-entropy alloys. *J Alloys Compd* 2015;624:270–8. DOI
6. Gludovatz B, Hohenwarter A, Catoor D, Chang EH, George EP, Ritchie RO. A fracture-resistant high-entropy alloy for cryogenic applications. *Science* 2014;345:1153–58. DOI
7. Senkov ON, Wilks GB, Miracle DB, Chuang CP, Liaw PK. Refractory high-entropy alloys. *Intermetallics* 2010;18:1758–65. DOI
8. Senkov ON, Wilks GB, Scott JM, Miracle DB. Mechanical properties of Nb<sub>25</sub>Mo<sub>25</sub>Ta<sub>25</sub>W<sub>25</sub> and V<sub>20</sub>Nb<sub>20</sub>Mo<sub>20</sub>Ta<sub>20</sub>W<sub>20</sub> refractory high entropy alloys. *Intermetallics* 2011;19:698–706. DOI
9. Zhang X, Chen Y, Hu J. Recent advances in the development of aerospace materials. *Prog Aerosp Sci* 2018;97:22–34. DOI
10. Steingrimsen B, Fan X, Feng R, Liaw PK. A physics-based machine-learning approach for modeling the temperature-dependent yield strengths of medium- or high-entropy alloys. *Appl Mater Today* 2023;31:101747. DOI
11. Senkov ON, Jensen JK, Pilchak AL, Miracle DB, Fraser HL. Compositional variation effects on the microstructure and properties of a refractory high-entropy superalloy AlMo<sub>0.5</sub>NbTa<sub>0.5</sub>TiZr. *Mater Design* 2018;139:498–511. DOI
12. Jensen JK, Welk BA, Williams REA, et al. Characterization of the microstructure of the compositionally complex alloy Al<sub>1</sub>Mo<sub>0.5</sub>Nb<sub>1</sub>Ta<sub>0.5</sub>Ti<sub>1</sub>Zr<sub>1</sub>. *Scripta Mater* 2016;121:1–4. DOI
13. Silver D, Huang A, Maddison CJ, et al. Mastering the game of Go with deep neural networks and tree search. *Nature* 2016;529:484–9. DOI
14. Huang W, Bai XM. Machine learning based on-the-fly kinetic Monte Carlo simulations of sluggish diffusion in Ni-Fe concentrated alloys. *J Alloys Compd* 2023;937:168457. DOI
15. Liu X, Xu P, Zhao J, Lu W, Li M, Wang G. Material machine learning for alloys: applications, challenges and perspectives. *J Alloys Compd* 2022;921:165984. DOI
16. Xiao L, Wang G, Long W, Liaw PK, Ren J. Fatigue life prediction of the FCC-based multi-principal element alloys via domain knowledge-based machine learning. *Eng Fract Mech* 2024;296:109860. DOI
17. Wei G, Byggmästar J, Cui J, Nordlund K, Ren J, Djurabekova F. Effects of lattice and mass mismatch on primary radiation damage in W-Ta and W-Mo binary alloys. *J Nucl Mater* 2023;583:154534. DOI
18. Zhu Y, Cui J, Guo X, Ren J. Multi-component thin films and coatings. *Mater Design* 2024;238:112664. DOI
19. Xiao L, Guo XX, Sun YT, et al. Sparse identification-assisted exploration of the atomic-scale deformation mechanism in multiphase CoCrFeNi high-entropy alloys. *Sci China Technol Sci* 2024;67:1124–32. DOI
20. Rickman JM, Chan HM, Harmer MP, et al. Materials informatics for the screening of multi-principal elements and high-entropy alloys. *Nat Commun* 2019;10:2618. DOI
21. Xiong J, Shi SQ, Zhang TY. Machine learning of phases and mechanical properties in complex concentrated alloys. *J Mater Sci Technol* 2021;87:133–42. DOI
22. Khatavkar N, Singh AK. Highly interpretable machine learning framework for prediction of mechanical properties of nickel based superalloys. *Phys Rev Mater* 2022;6:123603. DOI
23. Couzinie JP, Senkov ON, Miracle DB, Dirras G. Comprehensive data compilation on the mechanical properties of refractory high-entropy alloys. *Data Brief* 2018;21:1622–41. DOI
24. Stepanov ND, Yurchenko NY, Zharebtsov SV, Tikhonovsky MA, Salishchev GA. Aging behavior of the HfNbTaTiZr high entropy alloy.



- Mater Lett* 2018;211:87–90. DOI
25. Song HQ, Tian FY, Wang DP. Thermodynamic properties of refractory high entropy alloys. *J Alloys Compd* 2016;682:773–7. DOI
  26. Lin CM, Juan CC, Chang CH, Tsai CW, Yeh JW. Effect of Al addition on mechanical properties and microstructure of refractory  $\text{Al}_x\text{HfNbTaTiZr}$  alloys. *J Alloys Compd* 2015;624:100–7. DOI
  27. Hou S, Sun M, Bai M, Lin D, Li Y, Liu W. A hybrid prediction frame for HEAs based on empirical knowledge and machine learning. *Acta Mater* 2022;228:117742. DOI
  28. Gschneidner K, Russell A, Pecharsky A, et al. A family of ductile intermetallic compounds. *Nature Materials* 2003;2:587–90. DOI
  29. Zhang X, Li W, Ma J, Li Y, Zhang X, Zhang X. Modeling the effects of grain boundary sliding and temperature on the yield strength of high strength steel. *J Alloys Compd* 2021;851:156747. DOI
  30. Takeuchi A, Inoue A. Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element. *Mater Trans* 2005;46:2817–29. DOI
  31. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: a new perspective. *Neurocomputing* 2018;300:70–9. DOI
  32. Lu W, Xiao R, Yang J, Li H, Zhang W. Data mining-aided materials discovery and optimization. *J Materiomics* 2017;3:191–201. DOI
  33. Zhang H, Fu H, He X, et al. Dramatically enhanced combination of ultimate tensile strength and electric conductivity of alloys via machine learning screening. *Acta Mater* 2020;200:803–10. DOI
  34. Liddle AR. Information criteria for astrophysical model selection. *Mon Not R Astron Soc* 2007;377:L74–8. DOI
  35. Akaike H. A new look at the statistical model identification. *IEEE T Automat Contr* 1974;19:716–23. DOI
  36. Wang C, Zhong W, Zhao JC. Insights on phase formation from thermodynamic calculations and machine learning of 2436 experimentally measured high entropy alloys. *J Alloys Compd* 2022;915:165173. DOI
  37. Li GN, Zheng Y, Liu JY, et al. An improved stacking ensemble learning-based sensor fault detection method for building energy systems using fault-discrimination information. *J Build Eng* 2021;43:102812. DOI
  38. Stefenon SF, Ribeiro MHD, Nied A, et al. Time series forecasting using ensemble learning methods for emergency prevention in hydroelectric power plants with dam. *Electr Pow Syst Res* 2022;202:107584. DOI
  39. Mishra A, Kompella L, Sanagavarapu LM, Varam S. Ensemble-based machine learning models for phase prediction in high entropy alloys. *Comput Mater Sci* 2022;210:111025. DOI
  40. Zhang YF, Ren W, Wang WL, et al. Interpretable hardness prediction of high-entropy alloys through ensemble learning. *J Alloys Compd* 2023;945:169329. DOI
  41. Chen Q, He Z, Zhao Y, et al. Stacking ensemble learning assisted design of Al-Nb-Ti-V-Zr lightweight high-entropy alloys with high hardness. *Mater Design* 2024;246:113363. DOI
  42. Jiang H, Zheng W, Dong Y. Sparse and robust estimation with ridge minimax concave penalty. *Inform Sci* 2021;571:154–74. DOI
  43. Jiang H, Tao C, Dong Y, Xiong R. Robust low-rank multiple kernel learning with compound regularization. *Eur J Oper Res* 2021;295:634–47. DOI
  44. Jiang H, Luo S, Dong Y. Simultaneous feature selection and clustering based on square root optimization. *Eur J Oper Res* 2021;289:214–31. DOI
  45. Simon R. Fundamentals of data mining in genomics and proteomics. New York: Springer; 2007. DOI
  46. An Z, Mao S, Yang T, et al. Spinodal-modulated solid solution delivers a strong and ductile refractory high-entropy alloy. *Mater Horiz* 2021;8:948–55. DOI
  47. Ding XY, Zheng HY, Zhang PP, Luo LM, Wu YC, Yao JH. Microstructure and mechanical properties of WTaVCrTi refractory high-entropy alloy by vacuum levitation melting for fusion applications. *J Mater Eng Perform* 2022;32:7869–78. DOI
  48. Huang R, Wang W, Li T, et al. A novel AlMoNbHfTi refractory high-entropy alloy with superior ductility. *J Alloys Compd* 2023;940:168821. DOI
  49. Jiang W, Wang X, Li S, Ma T, Wang Y, Zhu D. A lightweight  $\text{Al}_{0.8}\text{Nb}_{0.5}\text{Ti}_2\text{V}_2\text{Zr}_{0.5}$  refractory high entropy alloy with high specific yield strength. *Mater Lett* 2022;328:133144. DOI
  50. Ma X, Hu Y, Wang K, et al. Microstructure and mechanical properties of a low activation cast WTaHfTiZr refractory high-entropy alloy. *China Foundry* 2022;19:489–94. DOI
  51. Goldstein A, Kapelner A, Bleich J, Pitkin E. Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. *J Comput Graphical Stat* 2015;24:44–65. DOI
  52. Apley DW, Zhu J. Visualizing the effects of predictor variables in black box supervised learning models. *J R Stat Soc B* 2020;82:1059–86. DOI
  53. Kusdhany MIM, Lyth SM. New insights into hydrogen uptake on porous carbon materials via explainable machine learning. *Carbon* 2021;179:190–201. DOI
  54. Rodríguez-Pérez R, Bajorath J. Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions. *J Comput Aided Mol Des* 2020;34:1013–26. DOI
  55. Ren J, Xiao L. Physics-guided neural network for fatigue life prediction of FCC-based multi-principal element alloys. *Scripta Mater* 2024;253:116307. DOI
  56. Lee C, Kim G, Chou Y, et al. Temperature dependence of elastic and plastic deformation behavior of a refractory high-entropy alloy. *Sci Adv* 2020;6:eaz4748. DOI
  57. Startt J, Kustas A, Pegues J, Yang P, Dingreville R. Compositional effects on the mechanical and thermal properties of MoNbTaTi refractory complex concentrated alloys. *Mater Design* 2022;213:110311. DOI
  58. Xue J, Shen B. Dung beetle optimizer: a new meta-heuristic algorithm for global optimization. *J Supercomput* 2022;79:7305–36. DOI



59. Giles SA, Sengupta D, Broderick SR, Rajan K. Machine-learning-based intelligent framework for discovering refractory high-entropy alloys with improved high-temperature yield strength. *npj Comput Mater* 2022;8:235. [DOI](#)