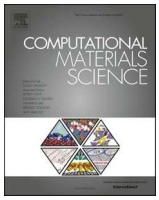
Computational Materials Science 239 (2024) 112976

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci



Full Length Article

Phase prediction of high-entropy alloys based on machine learning and an improved information fusion approach

Cun Chen [a,*], Xiaoli Han [a], Yong Zhang [b], Peter K. Liaw [c], Jingli Ren [a]

[a] School of Mathematics and Statistics, Henan Academy of Big Data, Zhengzhou University, Zhengzhou 450001, China
[b] State Key Laboratory for Advanced Metals and Materials, University of Science and Technology Beijing, Beijing 100083, China
[c] Department of Materials Science and Engineering, The University of Tennessee, Knoxville, TN 37996, USA

ARTICLE INFO

Keywords:
High-entropy alloys
Phase prediction
Machine learning
Conditional generative adversarial networks
Sparrow search algorithm
Information fusion

ABSTRACT

The phase design of high entropy alloys (HEAs) is an important issue since the phase structure affects the comprehensive properties of HEAs. Accurate prediction of phase classification can accelerate material design. In this paper, a new phase prediction framework is proposed using machine learning (ML) and an improved information fusion approach based on the Dempster-Shafer (DS) evidence theory. Considering that the classification results of different ML algorithms may conflict, and the traditional DS evidence theory cannot solve the problem of high conflict, we propose an improved information fusion approach based on the DS evidence theory. The basic probability assignment function is constructed using the ML algorithms. 761 HEAs samples are collected consisting of amorphous phase (AM), solid solution (SS), intermetallic compound (IM), and a mixture of SS and IM (SS + IM). For the small dataset of HEAs, we use a conditional generative adversarial network (CGAN) for data augmentation. Based on the enhanced dataset, the ML model is optimized by sparrow search algorithm (SSA), which can accelerate searching speed of model hyperparameters and improve the performance of the model. The results show that the proposed information fusion method performs better than several other existing techniques on the test set, and the prediction accuracy reaches 94.78 %. Meanwhile, the prediction accuracy of the proposed method is higher than that of the existing technology (93.17 %). It is proved that the proposed method can solve the high conflict problem effectively. Moreover, we present the interpretability analysis of the features by the Shapley additive explanations (SHAP) and the sensitivity matrix. A smaller atomic size difference δ (<6.6 %) is conducive to the formation of SS phase, while a larger δ (>6.6 %) is conducive to the formation of AM phase. A smaller enthalpy of mixing ΔH_{mix} tends to form AM phase. In binary and ternary alloy systems, IM phase can be extracted by the mixing enthalpy $\Delta S_{\mathrm{mix}} < 10$. In addition, we find that mean bulk modulus (K) and standard deviation of melting temperature (σ_T) are critical features to distinguish between SS and SS + IM.

1. Introduction

High-entropy alloys (HEAs) are refers to the composition containing at least five metallic elements, and the atomic percentage of each element is 5–35 atomic percent (at.%) [1,2]. Due to the excellent mechanical properties, HEAs have a broad application prospect [3–8]. The crystal structures exhibit different thermodynamic properties, such as high strength, great corrosion resistance and high temperature, high temperature softening resistance and good wear resistance. In general, the SS phase is closely related to the hardness and ductility of the HEAs [9], while the AM phase is related to the elasticity and electrical corrosion [10]. HEAs with intermetallic compound phases are generally better functional materials [11]. The formation of phase can significantly affect its microstructure, which in turn affects its physical and mechanical properties. Therefore, the prediction of phase information is crucial for the development of new HEAs.

The problem of phase design has aroused wide concern [12–14]. Several computational simulation methods are applied to the phase design of HEAs, such as first-principle [15,16], Monte Carlo simulation [17–19], molecular dynamics simulation [20,21], and calculation of phase diagrams (CALPHAD) [22,23]. However, the computational simulation methods are based on certain hypothetical framework, which are only applied for simple cases. For complex phase structures in HEAs, computational simulation methods may not be applicable. We tend to

* Corresponding author.
E-mail address: chencun@zzu.edu.cn (C. Chen).

https://doi.org/10.1016/j.commatsci.2024.112976
Received 20 December 2023; Received in revised form 20 March 2024; Accepted 21 March 2024
Available online 29 March 2024

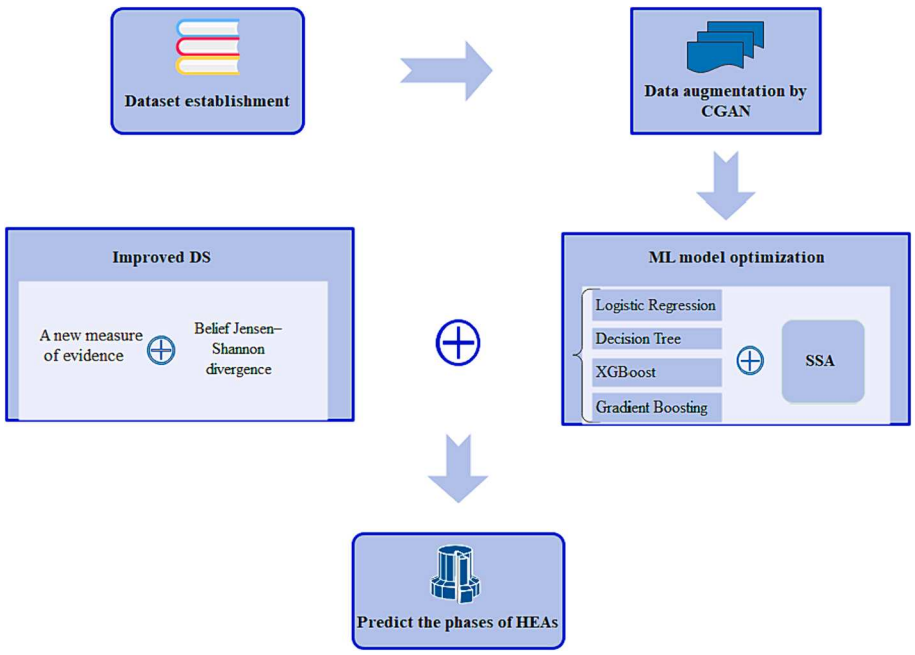


Fig. 1. Schematic flow of phase prediction based on CGAN and the improved DS.

develop data-driven phase design approach for HEAs based on real experimental data.

Machine learning (ML) can explore potential relationship between the features and the targets from real data, which has been used in the phase prediction of HEAs [24–28]. Many scholars have studied the phase classification by ML and obtained satisfactory results in two or three types of phase prediction [29–33]. For example, Krishna et al. applied ML algorithms to predict SS and SS + IM in a dataset containing 636 alloys, and the accuracy of the artificial neural network was up to 80 % [31]. Zhu et al. proposed a deep neural network structure using residual network to predict AM, SS, and IM phases, the accuracy of phase prediction was 81.9 % [33]. However, the prediction of multi-class of HEAs needs to be improved. For example, Lee et al. used generative adversarial networks to generate additional HEAs samples from 989 samples and constructed regularized deep neural networks for predicting the SS, SS + IM, AM and IM HEAs phases. The enhancement of the generative model significantly improved the performance of the model, and the prediction accuracy reached 93.17 % [34]. Qu et al. based several KNN models with k-dimensional tree algorithm and weighted voting to classify BCC, FCC, others, hexagonal close-packed (HCP) and IM using a dataset containing 2043 pieces of data, and the test accuracy is 93 % [35]. Singh et al. applied ML algorithms to predict the FCC, BCC, FCC + BCC, and intermetallic mixture (MIP) phases of HEAs based on 1200 original samples. The average accuracy is 92 % [36]. For the problem that the prediction results of multiple ML algorithms may be inconsistent, Hou et al. proposed a hybrid prediction frame for HEAs that combined empirical knowledge and Dempster-Shafer (DS) evidence theory. The model is verified by the prediction of SS, SS + IM, AM and IM in a dataset of 426 HEA samples. The results show that the phase prediction of quinaries, senaries and septenaries material systems are 87.8 ± 2.1 %, 86.7 ± 1.7 % and 83.3 ± 1.4 %, respectively [37]. The learning ability of a single ML algorithm is limited, while multiple ML algorithms predictions may conflict. In order to solve this problem, we apply DS evidence theory to predict the four phases of HEAs based on four ML models. The traditional DS evidence theory cannot solve the high conflict problem effectively [38–41]. Therefore, we propose an improved DS combination rule, which can effectively solve the problem of inconsistent prediction results of various algorithms and improve the prediction accuracy of multiphase classification problems based on small dataset.

In the present work, we collected 761 HEAs samples consisting of amorphous (AM), solid solution (SS), intermetallic compound (IM), and a mixture of SS and IM (SS + IM). Considering the small dataset, the conditional generative adversarial network (CGAN) is introduced for data augmentation. Sparrow search algorithm (SSA) is a powerful tool to optimize hyperparameters of ML algorithm and improve model performance [42]. We used SSA to optimize the hyperparameter of the basic classification models, such as Logistic Regression (LR), Decision tree (DT), eXtreme Gradient Boosting (XGBoost) and Gradient Boosting Decision Tree (GBDT). This paper proposes a phase prediction framework based on CGAN and an improved DS evidence theory which fuses ML models optimized by the SSA, which can improve the classification performance. The proposed model can achieve an accuracy of 94.78 % for phase classification, which performs better than several existing models. The advancement of the current work relates to three factors: (1) Considering the small dataset, CGAN is used to data augmentation (2) In the ML classification models, SSA is applied for hyperparameter optimization. It can accelerate searching speed of the hyperparameters and avoid falling into the local optimal solution. (3) The improved DS evidence theory can effectively solve high conflict problem and improve the accuracy of model prediction.

2. Materials and methods

2.1. Design strategy

The prediction frame of the improved DS evidence fusion method for HEAs phase is shown in Fig. 1. First, the dataset is established containing phases of HEAs: SS, AM, IM, and SS + IM. Second, CGAN is used to generate virtual training samples for each class of phases. Inception Score (IS) is applied to measure the quality of generated data. In addition, four ML algorithms are optimized by SSA in order to improve the performance of the model. The output of the optimized ML models is treated as mass functions. Finally, we propose an improved DS evidence theory, which can solve the confliction of results among different algorithms. The proposed information fusion method performs better than four ML algorithms optimized by SSA and traditional DS evidence theory on the test set.

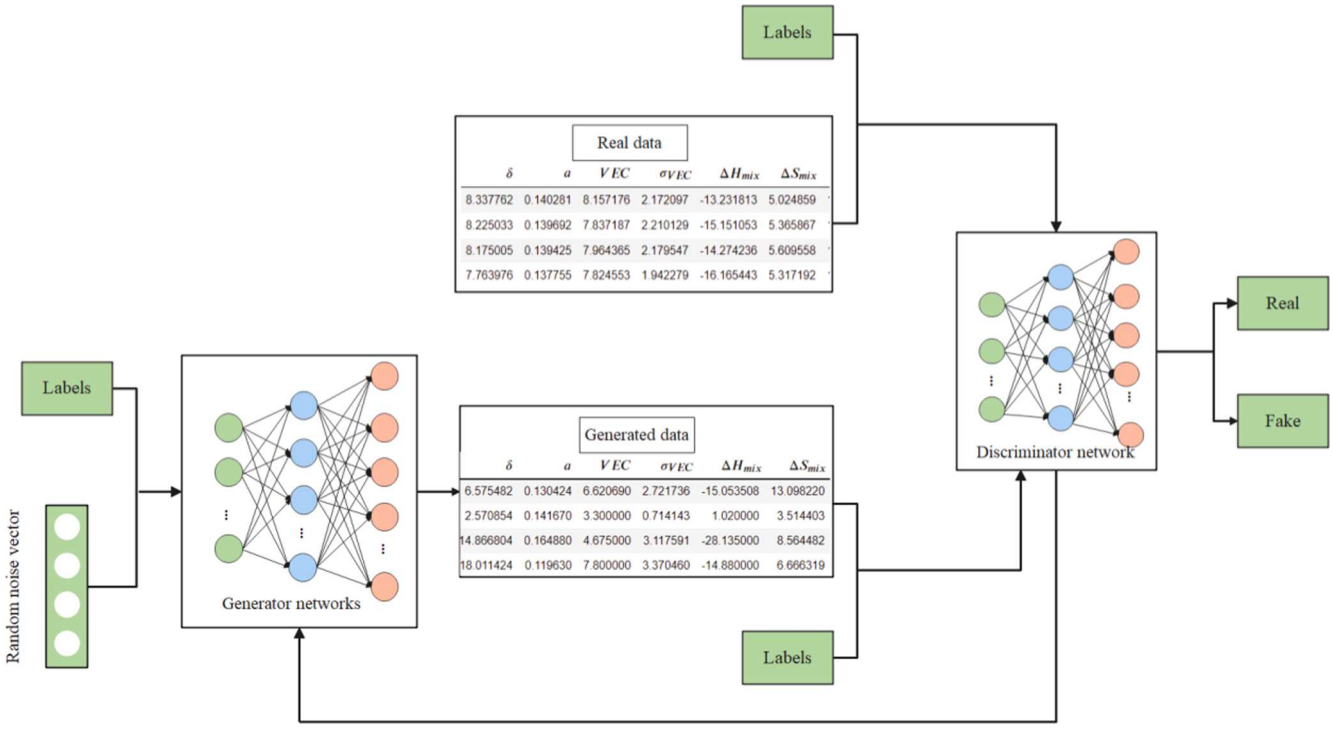


Fig. 2. The structure of CGAN model.

2.2. Conditional generative adversarial network

CGAN is a generative adversarial network (GAN) with constraints proposed by Mehdi Mirza in 2014 [43]. By adding additional conditional information, namely category label to the generator G and discriminator D of the original GAN, the conditional generation model is implemented. In addition to the prior input noise $P_z(z)$, the input of the generator has the conditional information y, namely the category label, so that we can generate the sample more precisely. The discriminator is the input sample labels and the corresponding features, which are combined to judge the authenticity of the sample. The objective function of CGAN is a two-player minimax game with conditional probability, and the loss function is described as follows,

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[\log D(x|y)] + E_{z\sim p_z(z)}[\log(1 - D(G(z|y)))]. \quad (1)$$

The schematic diagram of CGAN is shown in Fig. 2. The purpose of the discriminator is to distinguish the input data is real data or generate synthetic data $G(z|y)$. The purpose of the generator is to generate the data distribution as close as possible to the real data distribution $P_{data}(x)$. We expect the discriminator to distinguish well between real data and generated data. $D(x|y)$ represents the probability of x under condition y, where $D(x|y)$ to be close to 1 and $D(G(z|y))$ to be close to 0, corresponding to $V(D,G)$ maximization. The generator tries to imitate the real data to deceive the discriminator. $D(G(z|y))$ is expected to be close to 1, which corresponds to $V(D,G)$ minimization. During the training process, both sides strive to optimize their network, thus forming a competitive confrontation until both sides reach a dynamic equilibrium.

2.3. ML models optimized by SSA

The SSA is a swarm optimization approach proposed by Xue et al. based on group intelligence, foraging and anti-predation behaviors of sparrows [44]. In the simulation of sparrow foraging, sparrows can be divided into discoverers and participants. The location of sparrows and their food is constantly changing due to the presence of predators. The discoverer is responsible for finding food in the population and providing feeding areas and directions for the entire sparrow population, while the participant follows the discoverer to obtain food. Therefore, SSA is introduced to optimize the hyperparameters of Logistic Regression, Decision tree, eXtreme Gradient Boosting and Gradient Boosting models respectively. The use of SSA optimization algorithm can accelerate searching speed of model hyperparameters and improve model prediction performance.

2.4. DS evidence theory

The DS evidence theory was firstly proposed by Dempster [45] in 1967. It is further developed by Shafer [46] as a theory of uncertainty reasoning which is also a generalization of Bayesian theory. The frame of discernment Ω is a non-empty and mutually exclusive set, indicted by $\Omega = \{E_1, E_2, \ldots, E_n\}$.$2^\Omega = \{\varnothing, \{E_1\}, \ldots, \{E_n\}, \{E_1, E_2\}, \ldots, \{E_1, E_2, \ldots, E_i\}, \ldots, \Omega\}$ as a power set consists of a complete subset of Ω. A mass function is a mapping m from 2^Ω to $[0,1]$, the mass function can be also called as the basic probability assignment (BPA), indicted by,

$$m : 2^\Omega \rightarrow [0,1], \quad (2)$$

which satisfies $m(\varnothing) = 0$ and $\sum_{A\in 2^\Omega} m(A) = 1$. If $m(A) > 0$,A is a focal element.

Two independent BPAs m_1 and m_1 on 2^Ω by their orthogonal sum $m = m_1 \oplus m_2$ represent Dempster's rule of combination, is defined as follows,

$$m(A) = \begin{cases} \dfrac{1}{1-K} \displaystyle\sum_{B\cap C = A \subseteq \Omega} m_1(B) m_2(C), \\ 0, \end{cases} \quad (3)$$

$$K = \sum_{B\cap C = \varnothing} m_1(B) m_2(C) \quad (4)$$

where $B \in 2^\Omega$ and $C \in 2^\Omega$,$K \in [0,1]$ is a conflict coefficient that represents the conflict degree between BPAs m_1 and m_2. Dempster's rule of combination only makes sense at $K < 1$. If K proaches 1, the BPAs m_1 and

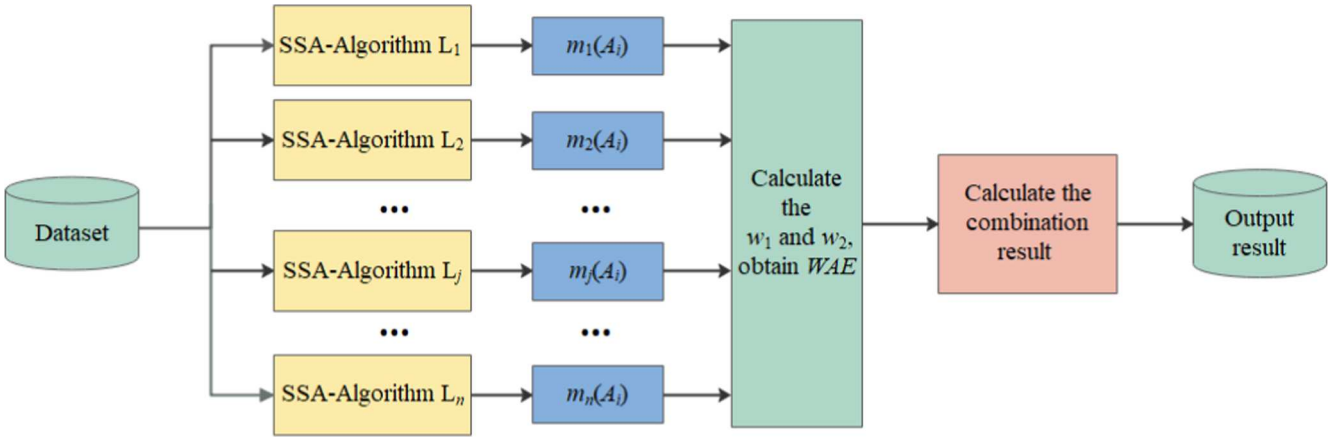


Fig. 3. Flowchart of the improved DS evidence fusion method.

m_2 are in high conflict. If K is closer to 0, the BPAs m_1 and m_2 are in low conflict. If $K = 1$, BPAs m_1 and m_2 are completely conflicting, i.e. the equation is mathematically meaningless, and traditional DS evidence theory may produce counterintuitive results.

2.5. Improved DS evidence theory

DS evidence theory is a powerful tool to handle with conflict problems. However, if there is a high conflict among several algorithms, the traditional DS evidence theory may produce counterintuitive results [47–50]. In this paper, we use the output of the ML models optimized by SSA to construct the basic belief assignments. An improved DS evidence fusion method based on a new measure of evidence and Belief Jensen–Shannon divergence is constructed to fuse the different algorithms. The flowchart for the improved approach is shown in Fig. 3.

Steps of the improved DS evidence fusion method are shown as follows:

Step 1: Build masses $M = \{m_1, m_2, \cdots m_i, \cdots m_n\}$ for each sample $x_p \in X = \{x_1, x_2, \cdots x_p, \cdots x_m\}$ by the ML algorithms $L = \{L_1, L_2, \cdots L_j, \cdots L_n\}$, where n represents the number of algorithms, m represents the total number of sample data.

Step 2: Calculate the distance $H(m_i)$ of each masse m_i as,

$$H(m_i) = \frac{1}{\sqrt{2}}\sqrt{\sum_{j=1}^{N}\frac{\left(\sqrt{m_i(A_j)} - \sqrt{c}\right)^2}{2^{|A_i|}-1}}, \tag{5}$$

where $c = 1/N$, N represents the total number of categories of phases;$m_i(A_j)$ represents the probability that sample x_p belongs to the j-th phase by i-th algorithm.

Step 3: The first weighting factor $w_1(i)$ of masse m_i is denoted as,

$$w_1(i) = \frac{[H(m_i)]^{-1}}{\sum_{k=1}^{n}[H(m_k)]^{-1}}. \tag{6}$$

Step 4: The divergence matrix is constructed as follows,

$$D = \begin{bmatrix} 0 & \cdots & BJS_{1i} & \cdots & BJS_{1n} \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ BJS_{i1} & \cdots & 0 & \cdots & BJS_{in} \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ BJS_{n1} & \cdots & BJS_{ni} & \cdots & 0 \end{bmatrix}, \tag{7}$$

where BJS_{ij} represents the divergence between two bodies of evidence $m_i, m_j (i, j = 1,2, \cdots n)$ is calculated by the Belief Jensen–Shannon divergence measure [51]. The Belief Jensen-Shannon divergence between the m_1 and m_2 is defined as,

$$BJS_{ij} = BJS(m_i, m_j) = \frac{1}{2}\left[S\left(m_i, \frac{m_i + m_j}{2}\right) + S\left(m_j, \frac{m_i + m_j}{2}\right)\right]$$

$$S(m_i, m_j) = \sum_k m_i(A_k)\log\frac{m_i(A_k)}{m_j(A_k)}, \sum_k m_i(A_k) = 1\left(k = 1, 2, \ldots, 2^{\Omega}; i = 1, 2\right)$$

Step 5: The second weighting factor $w_2(i)$ of masse m_i is denoted as,

$$w_2(i) = \frac{1/\overline{D}_i}{\sum_{k=1}^{n} 1/\overline{D}_k}, \tag{8}$$

where $\overline{D}_i = \frac{\sum_{j=1, i\neq j}^{n} BJS_{ij}}{n-1}$ represents the average distance of the masse m_i.

Step 6: Calculate the weighted average evidence WAE as follows,

$$WAE = \sum_{i=1}^{n}\left(w_1(i) \times w_2(i) \times m_i\right). \tag{9}$$

Step 7: Calculate the combination result of WAE as follows,

$$\widetilde{m}_p = \left(\left((WAE \oplus WAE)_1 \oplus WAE\right)_2 \oplus \cdots \oplus WAE\right)_{n-1}, \tag{10}$$

where $\oplus$ represents traditional Dempster's rule of combination [52,53].

Step 8: The element of matrix $\widetilde{m}_p$ represents the credibility of the phase of sample x_p, and the phase with the greatest credibility is taken as the output of the model.

3. Results and discussion

3.1. Dataset and features

We collected as-cast samples from previous literature [29,54,55,56,57] to develop phase prediction model for HEAs. Considering the lack of data of thermal history, we only consider compositions as the inputs to the machine learning models. 761 samples are collected with 12 candidate parameters, including atomic size difference δ, mean atom radius a, valence electron concentration VEC, standard deviations of valence electron concentration σ_{VEC}, electronegativity χ, standard deviation of electronegativity $\Delta\chi$, mean bulk modulus K, standard deviation of bulk modulus σ_K, average melting temperature T_m, standard deviation of melting temperature σ_T, mixing enthalpy ΔH_{mix} and mixing entropy ΔS_{mix}. This dataset is divided into 165 AM phases, 248 IM phases, 131 SS phases and 217 SS + IM phases, which refers to 56 elements (Al, B, In, Co, Cr, Cu, Fe, Mn, Ni, Pd, V, Mo, Ta, Hf, Zr, Nb, Sc, Y, Zn, Li, Mg, Sn, Ca, Gd, etc).

VEC, δ, $\Delta\chi$, ΔH_{mix} and ΔS_{mix} are the existing empirical design parameters, and on this basis, we add the 7 parameters namely a, σ_{VEC}, σ_K, σ_T, T_m, K and χ. These additional parameters were verified to have a positive effect on HEAs phase prediction [55]. The parameters of the

Table 1
The calculation formula of the parameters for HEAs phase.

Configurational parameters	Formulas
Mean atom radius	$a = \sum_{i=1}^{n} c_i r_i$
Atomic size difference	$\delta = \sqrt{\sum_{i=1}^{n} c_i \left(1 - \frac{r_i}{a}\right)^2}$
Mixing entropy	$\Delta S_{mix} = -R\sum_{i=1}^{n} c_i \ln c_i$
Enthalpy of mixing	$\Delta H_{mix} = \sum_{i=1, i<j}^{n} 4H_{ij} c_i c_j$
Average valence electron concentration	$VEC = \sum_{i=1}^{n} c_i VEC_i$
Standard deviation of *VEC*	$\sigma_{VEC} = \sqrt{\sum_{i=1}^{n} c_i (VEC_i - VEC)^2}$
Electronegativity	$\chi = \sum_{i=1}^{n} c_i \chi_i$
Standard deviation of electronegativity	$\Delta\chi = \sqrt{\sum_{i=1}^{n} c_i (\chi_i - \chi)^2}$
Mean bulk modulus	$K = \sum_{i=1}^{n} c_i K_i$
Standard deviation of bulk modulus	$\sigma_K = \sqrt{\sum_{i=1}^{n} c_i (K_i - K)^2}$
Average melting temperature	$T_m = \sum_{i=1}^{n} c_i T_{mi}$
Standard deviation of melting temperature	$\sigma_T = \sqrt{\sum_{i=1}^{n} c_i \left(1 - \frac{T_i}{T_m}\right)^2}$

samples are calculated according to the parameter formulas in Table 1. c_i, r_i, χ_i, K_i and T_{mi} are the percentage, atomic radius, pauling electronegativity, bulk modulus and melting point of the i-th element of the alloy, and H_{ij} is the heat of mixing between the i-th and j-th elements in Table 1. In this paper, all the codes are run in Python 3.8.5. The Pearson correlation coefficient is used to describe the correlation between each two parameters. As shown in the Fig. 4, the Pearson correlation coefficient ranges from −0.84 to 0.75, with no strong correlation between any two design parameters. Due to the large number of parameters, we choose the t-stochastic neighbor embedding (t-SNE) method for dimension reduction after data normalization to explore the structure of the data. The t-SNE constructs a Gaussian distribution and a T-distribution in a high-dimensional data space and a low-dimensional embedding space, respectively. The dimensionality reduction of data is achieved as shown in Fig. 5, and the entire dataset containing four types of HEAs phases is distributed in a two-dimensional feature space. It is obvious that the SS and SS + IM phases are indistinguishable from each other, which indicates that their original features are similar. Some samples of the AM and IM phase are located near the locations of the SS and SS + IM phases. In general, the classification of the four types of phases is not a linear problem. In the following, we choose DS evidence theory to solve the classification prediction by fusing the results of different ML models.

3.2. Data augmentation

The training set comprises 85 % of the dataset and contains 646 HEAs samples, of which 148 AM phases, 203 IM phases, 113 SS phases, and 182 SS + IM phases. The CGAN is employed to generate 100 data for each of the four types of phases. Both the generator and discriminator are fully connected networks and the activation function is LeakyReLU. The generator has two hidden layers and the discriminator has one hidden layer, which is to prevent the discriminator from being too strong. BatchNormalization and Dropout are added into the generating network, which can generate better HEAs data.

To verify the quality of the generated data, we chose the Inception Score (IS) proposed by Salimans et al. in 2016 [58]. IS is a comprehensive GAN evaluation index in terms of the quality and diversity of the generated data. The formula for calculating the IS of the generator is:

$$IS(G) = \exp\left[\sum_i p(x_i) D_{KL}(P(y|x_i)|p(y))\right], \tag{11}$$

where $p(x_i) = 1/N, N = 4; P(y|x_i)$ represents the probability that the i-th generated sample belongs to $y; p(y)$ represents the probability that all the

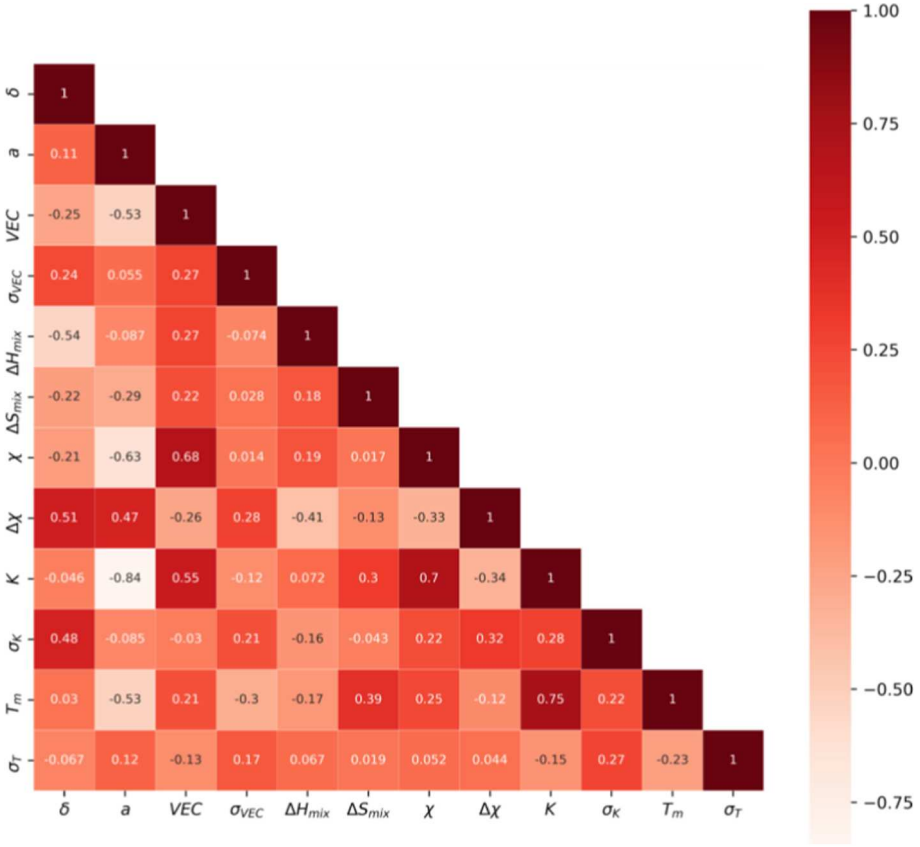


Fig. 4. Pearson correlation coefficients of all the features.

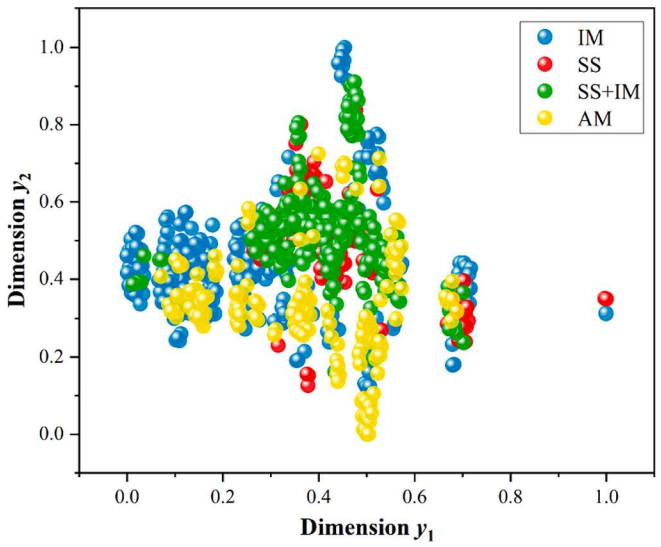


Fig. 5. The result of dimensionality reduction by t-SNE.

Table 2
The IS evaluation results of the CGAN.

$H(y\|x)$	$H(y)$	IS
0.1111	1.3857	3.5773

Table 3
Prediction accuracy of ML models and SSA-ML algorithms.

Method	LR	DT	XGBoost	GBDT
Basic ML	83.45 %	79.13 %	86.96 %	86.96 %
SSA-ML	86.09 %	88.70 %	93.04 %	93.04 %

generated samples belong to y;$D_{KL}(p(y|x_i)|p(y))$ is Kullback-Leibler divergence, which measures the similarity of two probability distributions [59].

IS can also be simply expressed as $IS(G) = e^{H(y)-H(y|x)}$.$H(y)$ is the information entropy of y, which means the uncertainty. A larger $H(y)$ means a higher diversity of the sample.$H(y|x)$ reflects the quality of the generated data, and a smaller $H(y|x)$ suggests that the generated data is closer to the real data. The values of IS are calculated in Table 2. The $H(y|x) = 0.1111$, which is close to the ideal result of 0, indicating that the quality of the generated data is good. The difference between $H(y) = 1.3857$ and the optimal result $\log K(\log 4 \approx 1.3863)$ is 0.0006, indicating that the generated data is of high diversity. In addition, IS = 3.5773 is close to 4, which means that the quality of the generated data is high.

3.3. ML models optimization

Four ML models including LR, DT, XGBoost, and GBDT are trained based on 646 original data and 400 data generated by CGAN respectively. To improve the performance of the base classifier, the relevant hyperparameter of the ML models are optimized using SSA. The regularization intensity 'C' of an important parameter of LR is optimized. For DT, three hyperparameters 'max_depth', 'max_features' and 'criterion', respectively are optimized. The hyperparameter 'objective' of XGBoost is set to 'multi: softproba', and we optimize 'n_estimators', 'max_depth', 'learning_rate' and 'gamma'. For GBDT, four important hyperparameters, including 'n_estimators', 'learning_rate', 'subsample' and 'max_depth' are selected to optimize.

To verify the advantages of the SSA algorithm, we conducted comparative experiments on LR, DT, XGBoost and GBDT algorithms with and without SSA optimization, respectively. The results of the experiment are listed in Table 3. The default values for the hyperparameters used in the ML models are sklearn defaults. The default hyperparameter setting of LR is as follows, 'penalty'='l2', 'C'=1.0. The default hyperparameter setting of DT is as follows, 'criterion'='gini', 'splitter'='best', 'max_depth'=None, 'max_features'=None. The default hyperparameter setting of XGBoost is as follows, 'n_estimators'=100, 'max_depth'=6, 'learning_rate'=0.3, 'gamma'=0. The default hyperparameter setting of GBDT is as follows, 'n_estimators'=100, 'subsample'=1.0, 'max_depth'=3, 'learning_rate'=0.1. LR has the lowest prediction accuracy, probably due to the fact that it cannot solve

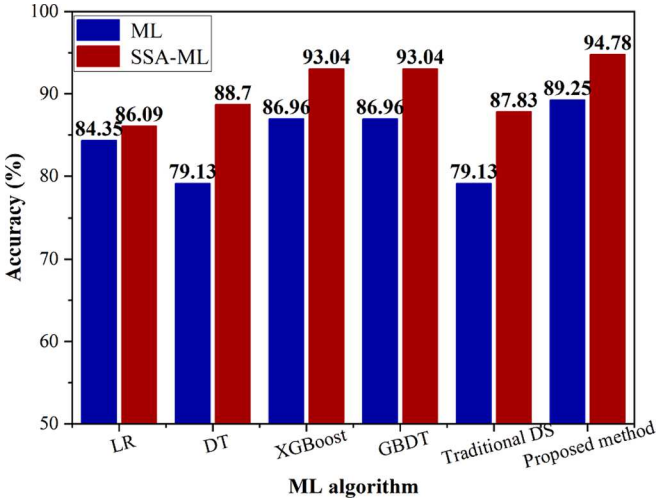


Fig. 6. Comparison of the proposed method with ML algorithms and traditional DS evidence theory.

Table 5
Comparison of different methods on various performances.

Algorithm	Precision	F1-Score	Recall
SSA-LR	89.66	87.19	86.09
SSA-DT	88.74	88.63	88.70
SSA-XGBoost	93.42	92.91	93.04
SSA-GBDT	93.03	92.90	93.04
SSA-DS	87.92	87.71	87.83
The proposed method	**94.78**	**94.75**	**94.78**

Table 4
rediction results of the SSA-ML algorithms and the proposed DS evidence theory.

Alloys	SSA-LR	SSA-DT	SSA-XGBoost	SSA-GBDT	DS	Proposed DS DSDSDS	True Phase
$Zr_{70}Pd_{30}$	AM	IM	AM	AM	IM	AM	AM
MoNbTaVW	SS	AM	SS	SS + IM	AM	SS	SS
AlCoCrCuFeNi	SS + IM	AM	SS + IM	SS + IM	AM	SS + IM	SS + IM
$AlCo_2CuFeNi$	SS + IM	SS + IM	SS + IM	SS	SS	SS + IM	SS + IM
$AlBe_4Mn$	IM	AM	IM	IM	AM	IM	IM
$Co_2Mo_{1.75}Ni_2VW_{1.75}$	SS	IM	SS + IM	SS + IM	IM	SS + IM	SS + IM
$NaZn_{13}$	IM	AM	IM	IM	AM	IM	IM

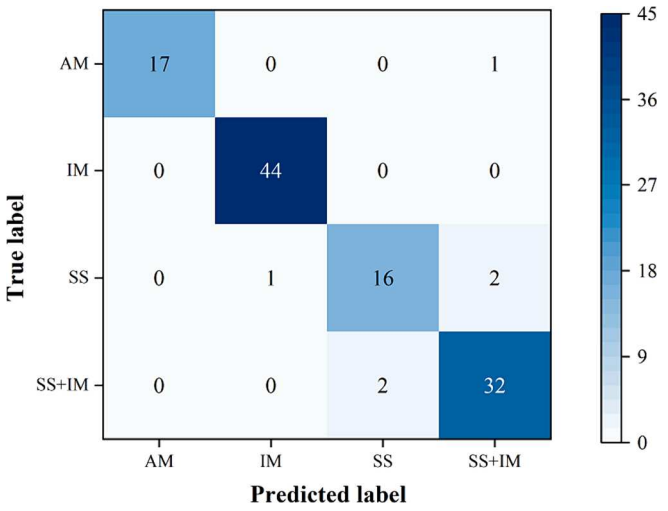


Fig. 7. Confusion matrix for phase prediction of the proposed method.

nonlinear problems. The performance of DT, XGBoost and GBDT optimized by SSA is significantly improved. The accuracy of DT increases significantly due to its hyperparameter 'max_depth' has a significant effect on the performance of the model. SSA can quickly find the global optimal value of 'max_depth' to improve the accuracy of the model. The hyperparameters 'n_estimators', 'max_depth' and 'learning_rate' of XGBoost and GBDT affect the training results of the models. SSA can help ML algorithms to optimize the hyperparameters, which makes the training results better than basic ML models.

3.4. ML models construction

In the following, we construct a phase prediction model using the proposed DS evidence theory to fuse the classification results of four SSA-ML algorithms including SSA-LR, SSA-DT, SSA-XGBoost and SSA-GBDT, which can solve the problem of high conflict among different algorithms.

In order to verify the high conflict resolution capability of the proposed DS evidence theory, we compare the output results of the proposed method and the SSA-ML algorithms. Several representative samples are listed in Table 4. For each alloy, the prediction results of different algorithms may not be consistent. For example, the true phase of $Co_2Mo_{1.75}Ni_2VW_{1.75}$ is SS + IM phase. SSA-XGBoost and SSA-GBDT predicted its phase correctly, while SSA-LR predicted it as SS phase and SSA-DT predicted it as IM phase. For alloys MoNbTaVW, although SSA-LR and SSA-XGBoost algorithms predicted its phase correctly, SSA-DT has a belief value of 1 for the AM phase. Therefore, there is a high conflict that it is difficult to make a correct judgment with the traditional DS evidence theory. In this situation, the improved DS evidence theory proposed in this paper can solve the high conflict problem effectively, and it can predict the phases more correctly.

To further demonstrate the effectiveness of the proposed method, we compared it with the performance of different ML algorithms and traditional DS evidence theory. Fig. 6 draws the accuracy of different ML algorithms optimized and not optimized by SSA and it shows the comparison of the accuracy of the proposed method with the traditional DS evidence theory and different SSA-ML algorithms. No matter whether the ML models are optimized by SSA or not, the traditional DS evidence theory cannot improve the prediction accuracy of the model, which is lower than the accuracy of the ensemble model XGBoost, due to the traditional DS evidence theory fails to solve high conflict problems effectively. As exhibited in Fig. 6, the accuracy of the proposed method reaches 94.78 %, which is higher than other methods. Moreover, we also compare other indicators of SSA-ML algorithms, such as precision, F-score and recall are shown in Table 5. The precision, F-score and recall of the proposed method are 94.78 %, 94.75 % and 94.78 %, respectively. In addition, to further study the phase prediction results of the proposed method, we present the confusion matrix. In Fig. 7, of the 115 test samples, 109 samples are predicted correctly. The accuracy of AM phase is 100 %, IM phase is 97.78 %, SS phase is 88.89 %, and SS + IM phase is 91.43 %. The above results show that the proposed method is better than four SSA-ML algorithms and the traditional DS evidence theory method. The reason is that the improved DS evidence theory assigns reasonable

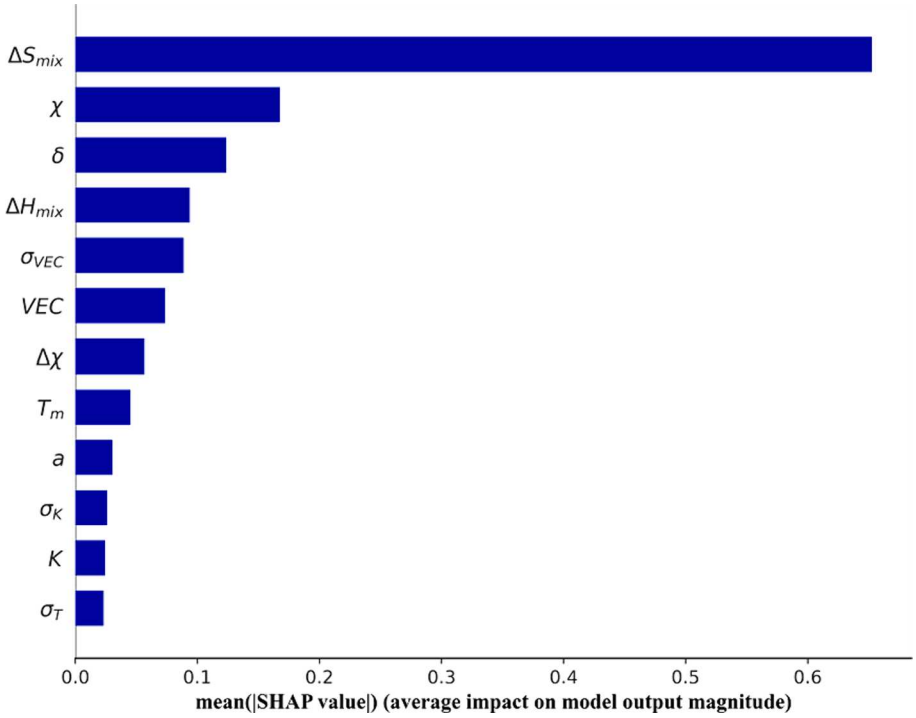


Fig. 8. Assessment of parameters importance level for the mean SHAP values.

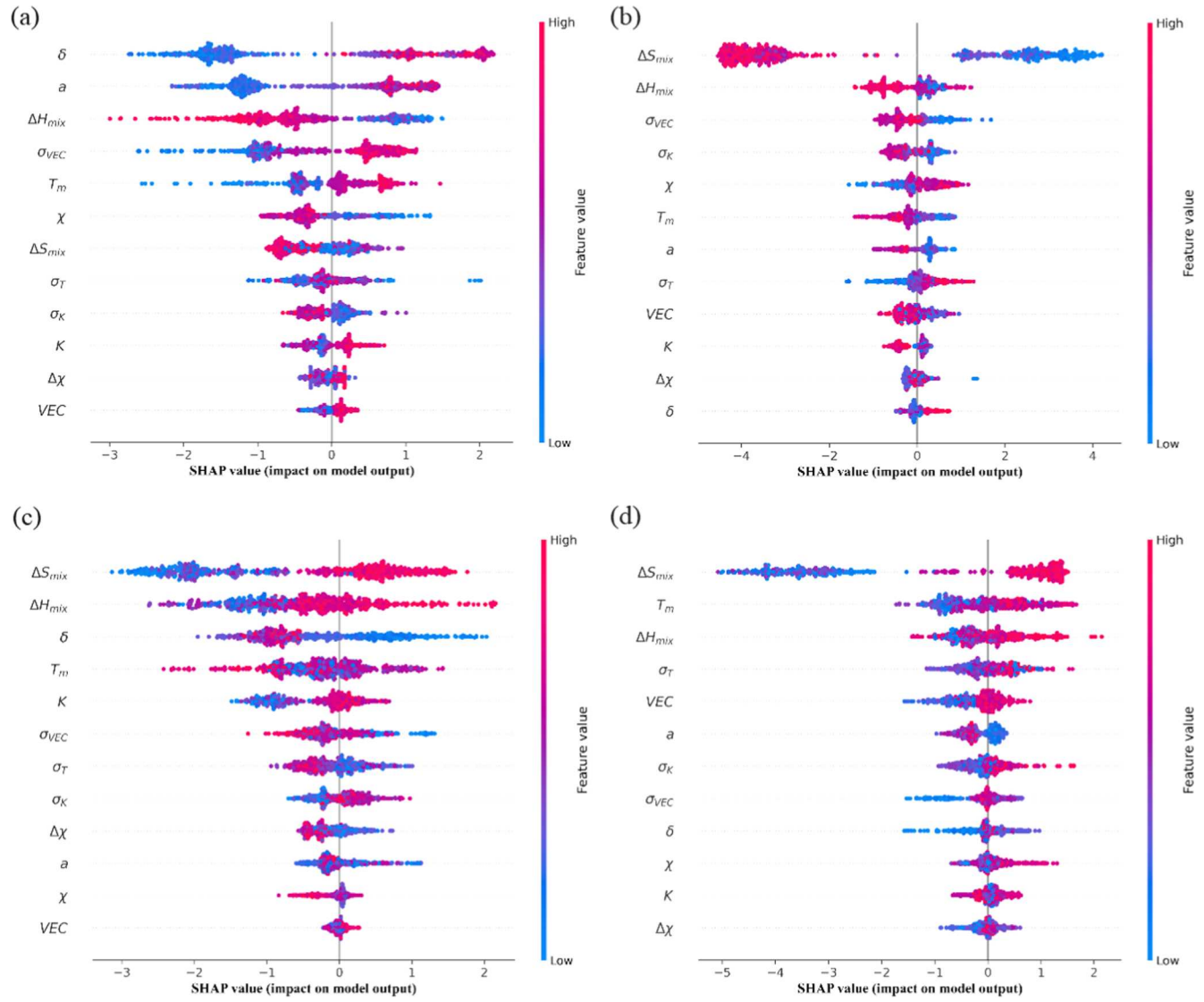


Fig. 9. Comparison of SHAP values of 12 design parameters. The SHAP value of different features for the prediction of (a) AM phase, (b) IM phase, (c) SS phase and (d) SS + IM phase.

weight to different algorithms, which can solve the problem of high conflict.

3.5. SHAP method and sensitivity matrix for interpretability analysis

In this part, the influence of features on the HEAs phase is analyzed by the SHapley Additive exPlanation (SHAP) [60] and the sensitivity matrix via artificial neural network (ANN) [55,61]. Firstly, the SHAP method is used for interpretability analysis. The importance of parameters can be obtained: $\Delta S_{\mathrm{mix}} > \chi > \delta > \Delta H_{\mathrm{mix}} > \sigma_{VEC} > VEC > \Delta\chi > T_{\mathrm{m}} > a > \sigma_K > K > \sigma_T$, which is shown in Fig. 8. In Fig. 9, each sample point represents an alloy sample, and the color represents the magnitude of the feature value. On the *y*-axis, features are ranked by importance from top to bottom. SHAP values represent the contribution of features to the output of the ML model. A positive SHAP value indicates that the feature increases the output of the ML model. On the contrary, a negative SHAP value indicates that the feature reduces the output of the ML model. The coverage of a feature is wider, the greater its effect on phase. As shown in Fig. 9, the larger δ value and smaller ΔH_{mix} produce the positive SHAP value, facilitating the formation of AM phase, which is reasonable and demonstrated in Fig. 11 (a). Large δ can cause serious lattice distortion, reduce the phase transformation rate and even lead to the formation of AM structures [62]. Negative ΔH_{mix} is a necessary condition for forming AM alloys. The stronger the interaction between atoms, the higher the possibility of forming an AM phase is [63]. ΔS_{mix} is the most important feature among the 12 features for the prediction of IM, SS and SS + IM phases (see Fig. 9(b)-(d)). Multi-component HEAs with equal atomic ratio have higher ΔS_{mix} than conventional alloys. According to $\Delta G_{mix} = \Delta H_{mix} - T\Delta S_{mix}$, high ΔS_{mix} reduces Gibbs energy ΔG_{mix}, which means a higher degree of confusion in alloy system. A higher ΔS_{mix} contributes to the formation of SS phase, which is more stable than other ordered phases and IM phase [64,65]. For the binary or ternary alloys with IM phase, the corresponding ΔS_{mix} value is small. It makes sense that a small ΔS_{mix} produces a positive SHPA value, which is conducive to the formation of IM phase. The smaller *VEC* value has a positive effect on the IM phase, while the larger *VEC* corresponds to a positive SHAP value in the prediction of SS and SS + IM phases, indicating that the larger *VEC* is conducive to produce the SS and SS + IM phases. Moreover, from Fig. 9, we find that there is an important influence for χ on the formation of AM and IM phases, while the influence of χ on the formation of SS and SS + IM phases is not significant.

In the following, the sensitivity matrix is introduced to further analyze the influence of parameters on the phase formation of HEAs. The sensitivity matrix is the product of all linear transformation matrices in

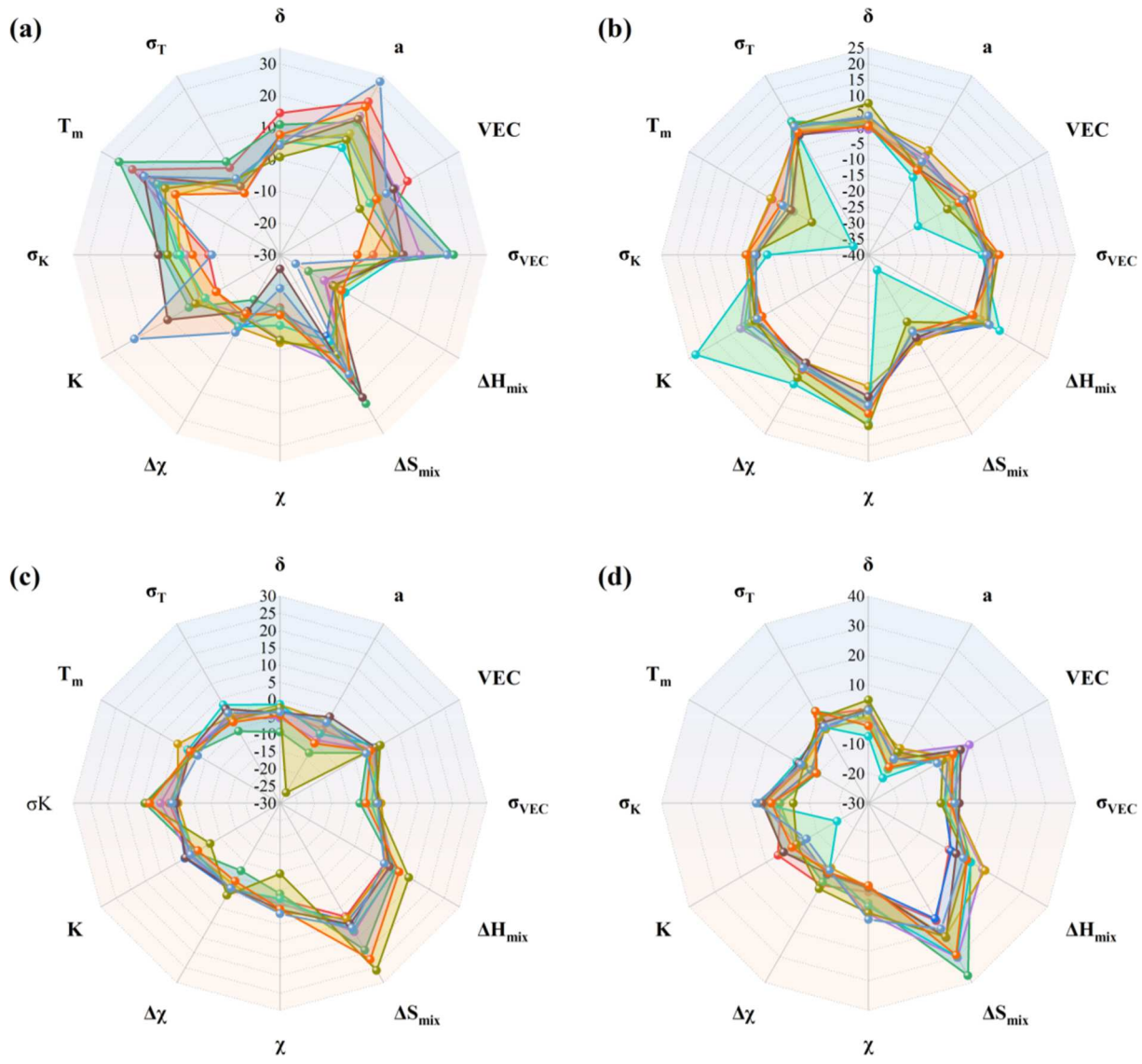


Fig. 10. The sensitivity values of 12 parameters for (a) AM phase, (b) IM phase, (c) SS phase and (d) SS + IM phase.

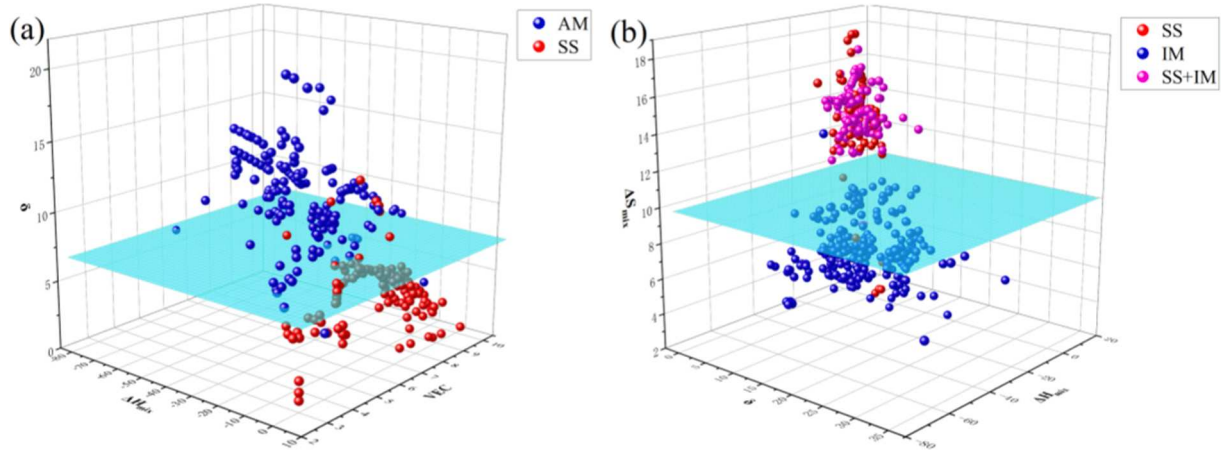


Fig. 11. Scatter plot of (a) design parameters ΔH_{mix}, *VEC* and δ for SS and AM phases, and (b) design parameters δ, ΔH_{mix} and ΔS_{mix} for SS, SS + IM and IM phases.

ANN, which can be used to assess the sensitivity of the features on the phases. The trained ANN in this paper contains an input layer, a hidden layer and an output layer. For each type of phase, ANN is trained ten times and the corresponding sensitivity matrix is calculated to obtain reliable results as shown in Fig. 10. A positive sensitivity value indicates that the feature have a positive correlation with the phase. On the

contrary, a negative sensitivity value indicates that the feature have a negative correlation with the phase. From Fig. 10 (a), it can be seen that ΔH_{mix}, a, ΔS_{mix}, χ, σ_{VEC} and T_m play an important role on the formation of AM phase. The feature importance is as follows, $a > T_m > \Delta H_{mix} > \chi > \Delta S_{mix}$. Negative sensitivity values of ΔH_{mix} and χ indicate that negative correlation with the AM phase. A positive σ_{VEC} is in favor of the formation of AM, while the σ_{VEC} shows a smaller negative correlation with IM, SS and SS + IM. The sensitivity value of $\Delta\chi$ suggests that it can promote the formation of IM, while restrain the formation of AM, SS and SS + IM. In addition, from Fig. 10 (c) and Fig. 10 (d), we find that low σ_{VEC} and σ_T are favorable for the formation of SS, while high σ_{VEC} and σ_T are beneficial to form SS + IM. As shown in Fig. 9 and Fig. 10, K has a large positive effect on SS, whereas it has a small effect on SS + IM. Lager σ_K favors the formation of AM, SS and SS + IM, while smaller σ_K favors the formation of IM. Based on the above analysis, K and σ_T are critical features to discriminate between SS and SS + IM.

In order to study the effect of parameters on phase formation, we designed a three-dimensional spatial coordinate system to visualize the data. According to the Fig. 11 (a), it is easy to form SS phase if the δ is less than 6.6 %, while it is easy to form AM phase if the δ is greater than 6.6 %. The negative ΔH_{mix} tends to form AM phase. The SS phases are mainly distributed in $\delta < 6.6$ % and $-18 < \Delta H_{mix} < 5$, which is consistent with the forming conditions of SS phase proposed by Guo et al [66]. In addition, IM phase can be well extracted by $\Delta S_{mix} < 10$ in Fig. 11 (b). The IM phase relatively concentrated in binary and ternary alloys, or HEAs with principal elements greater than 6 [67]. In this paper, for binary and ternary alloys, the IM phase is easy to form if ΔS_{mix} is less than 10; For alloys with principal elements greater than 3, the SS and SS + IM phases are easy to form if ΔS_{mix} is greater than 10.

4. Conclusions

In this study, we propose a new information fusion rule to predict the phases of HEAs based on machine learning models. The improved DS evidence theory combines a new measure of evidence and Belief Jensen–Shannon divergence, which can solve the high conflict issues effectively. The basic probability assignment function is constructed by the outputs ML algorithms. 761 HEAs samples are collected with 12 features. Considering the shortage of the raw data, CGAN is used to generate 100 samples for each phase. Based on the augmented dataset, the hyperparameters of four ML algorithms LR, DT, XGBoost, and GBDT are optimized using SSA, which can improve the performance of the model. To verify the validity of the proposed method, it is compared with the four SSA-ML models and the traditional DS evidence theory. The result shows that the proposed method performs best with an accuracy of 94.78 %. The reason is that the improved DS evidence theory can solve the problem of high conflict among the predicted results of different algorithms. In addition, explainable analysis of the ML model is presented. We find that the parameters ΔS_{mix}, χ, δ, ΔH_{mix}, σ_{VEC} and *VEC* have the greatest influence on formation of HEAs phase. In addition, the SS and AM phases can be distinguished well by $\delta = 6.6$ %. The binary and ternary alloys tend to form the IM phase if ΔS_{mix} is less than 10; The alloys with principal elements greater than 3 tend to form SS and SS + IM phases if ΔS_{mix} is greater than 10. K and σ_T are critical features to distinguish between SS and SS + IM. The method proposed in this paper has wide application prospect. It is not only suitable for phase prediction of HEAs but also can be applied in ceramic material, magnetic material, polymeric materials, etc.

CRediT authorship contribution statement

Cun Chen: Writing – review & editing, Supervision, Methodology, Conceptualization. **Xiaoli Han:** Writing – original draft, Software, Data curation. **Yong Zhang:** Writing – review & editing. **Peter K. Liaw:** Writing – review & editing, Resources. **Jingli Ren:** Writing – review & editing, Supervision, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

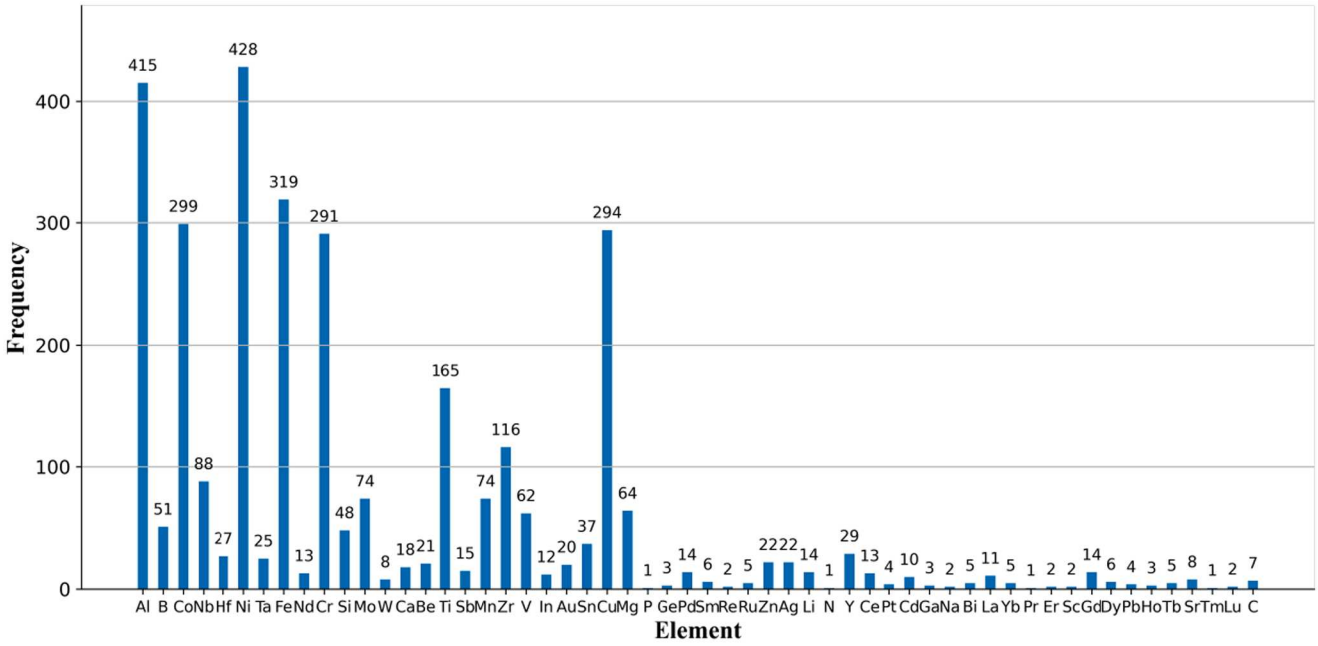
I have shared the link to my code in the data availability statement

Acknowledgements

This work was supported by the NSFC grant (No. 52071298); the National Science Foundation (DMR–1611180, 1809640, and 2226508); and the Army Research Office, United States (W911NF-13–1-0438 and W911NF-19–2-0049).

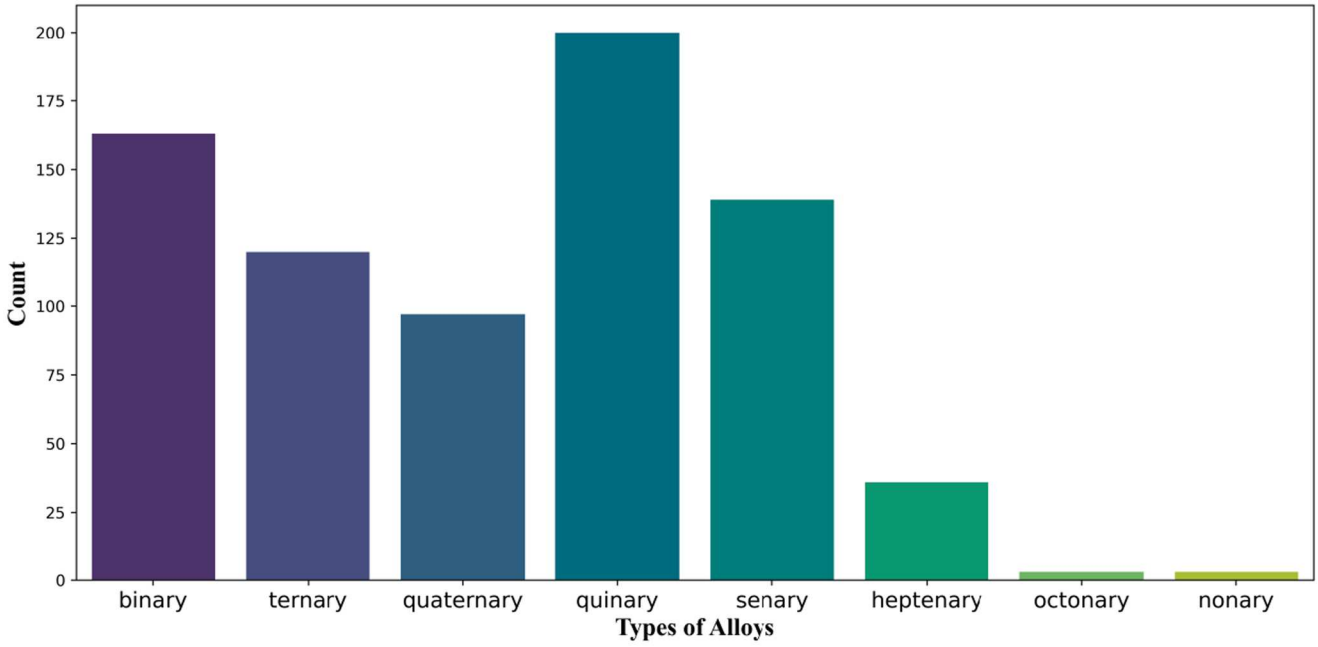
Appendix A. . Details regarding of alloys

Fig. 1: The distribution of the different elements.



From Fig. 1, it can be seen that the occurrence frequency of Ni elements is 428 times. Al, Fe, Co, Cu and Cr elements appear more than 290 times. P, N, Pr and Tm elements only appear once.

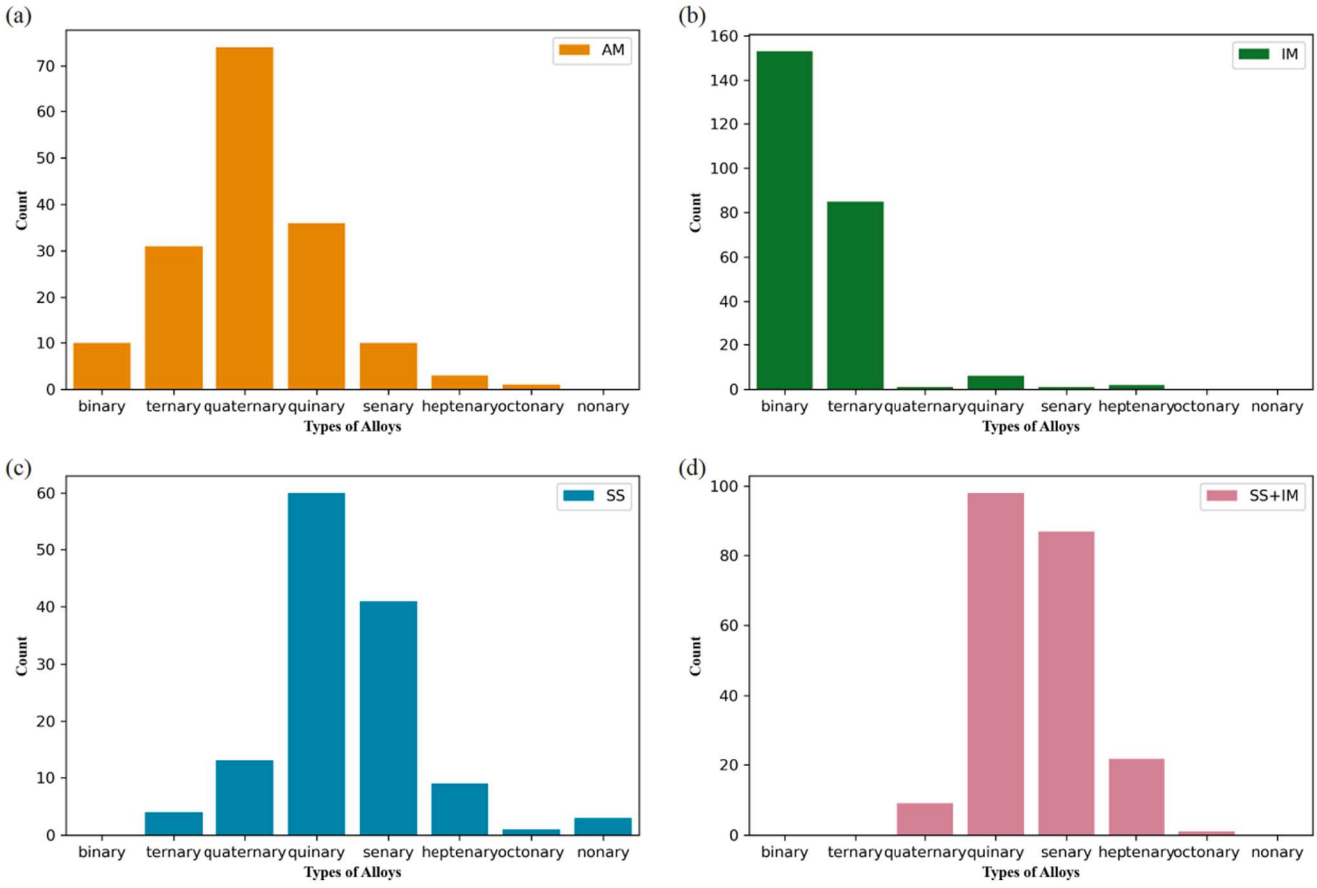
Fig. 2: The number of alloys with different components



.

It can be seen from Fig. 2, the number of binary, ternary, quinary and senary alloys exceeds 100, while the number of octonary and nonary alloys are less than 10.

Fig. 3: The distribution of different phases in different alloy systems. (a) AM, (b) IM, (c) SS, and (d) SS + IM.



The distribution of different phases in the alloy is further analyzed. As can be seen from Fig. 3, AM phase is mainly distributed in ternary, quaternary and quinary alloys. IM phase is mainly distributed in binary and ternary alloys. SS phase is mainly distributed in quinary and senary alloys. SS + IM phase is mainly distributed in quinary, senary and heptenary alloys. In Table1, it shows the minimum and maximum values of the mass percentage for each chemical element.

Table 1. The minimum and maximum values of the mass percentage for each chemical element.

Element	minimum	maximum
Ag	0.044928162	0.888843111
Al	0.008585424	0.769377008
Au	0.240626947	0.937238295
B	0.001960166	0.827823867
Be	0.011312497	0.615362983
Bi	0.301799417	0.851433111
C	0.004735039	0.066609972
Ca	0.045032696	0.798446138
Cd	0.086501835	0.932759861
Ce	0.215262144	0.867953914
Co	0.030891735	0.915986632
Cr	0.047113387	0.905828034
Cu	0.012552243	0.887991504
Dy	0.197549175	0.718867507
Er	0.724665391	0.925366528
Fe	0.009990466	0.905135288
Ga	0.177255739	0.492257002
Gd	0.191166817	0.828244978
Ge	0.275865998	0.484006397
Hf	0.07343273	0.891959422
Ho	0.225154262	0.609833527
In	0.110572895	0.685015412
La	0.061315202	0.84757756
Li	0.010021621	0.366592323
Lu	0.212708794	0.235630791
Mg	0.028721909	0.702231058

(continued on next page)

(*continued*)

Element	minimum	maximum
Mn	0.078082474	0.853201943
Mo	0.0394614	0.911090713
N	0.017808794	0.017808794
Na	0.026336591	0.091438799
Nb	0.034153496	0.774637319
Nd	0.266817118	0.84242495
Ni	0.030691594	0.942155394
P	0.078582122	0.078582122
Pb	0.344678444	0.759373053
Pd	0.01911286	0.986298029
Pr	0.798937849	0.798937849
Pt	0.783327979	0.955930875
Re	0.46313983	0.932448673
Ru	0.251380391	0.71407376
Sb	0.138860695	0.818609654
Sc	0.151736055	0.276099239
Si	0.002166211	0.524458551
Sm	0.191129924	0.755653835
Sn	0.019640284	0.913498165
Sr	0.193004097	0.540722704
Ta	0.199574671	0.692605068
Tb	0.193209171	0.613510905
Ti	0.02885021	0.726513887
Tm	0.205366043	0.205366043
V	0.023072951	0.5645781
W	0.215736666	0.820293153
Y	0.05218703	0.548942371
Yb	0.381162166	0.681313489
Zn	0.076876044	0.973663409
Zr	0.049622473	0.83529699

References

[1] J.W. Yeh, S.K. Chen, S.J. Lin, J.Y. Gan, T.S. Chin, T.T. Shun, C.H. Tsau, S.Y. Chang, Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes, Adv. Eng. Mater. 6 (5) (2004) 299–303, https://doi.org/10.1002/adem.200300567.

[2] B. Ren, Z.X. Liu, D.M. Li, L. Shi, B. Cai, M.X. Wang, Effect of elemental interaction on microstructure of CuCrFeNiMn high entropy alloy system, J. Alloys Compd. 493 (2010) 148–153, https://doi.org/10.1016/j.jallcom.2009.12.183.

[3] C. Wen, Y. Zhang, C. Wang, D. Xue, Y. Bai, S. Antonov, L. Dai, T. Lookman, Y. Su, Machine learning assisted design of high entropy alloys with desired property, Acta Mater. 170 (2019) 109–117, https://doi.org/10.1016/j.actamat.2019.03.010.

[4] N. Hua, W. Wang, Q. Wang, Y. Ye, S. Lin, L. Zhang, Q. Guo, J. Brechtl, P.K. Liaw, Mechanical, corrosion, and wear properties of biomedical Ti-Zr-Nb-Ta-Mo high entropy alloys, J. Alloys Compd. 861 (2021) 157997, https://doi.org/10.1016/j.jallcom.2020.157997.

[5] C. Wen, C. Wang, Y. Zhang, S. Antonov, D. Xuee, T. Lookmanf, Y. Su, Modeling solid solution strengthening in high entropy alloys using machine learning, Acta Mater. 212 (2021) 116917, https://doi.org/10.1016/j.actamat.2021.116917.

[6] W. Wang, K. Yang, Q. Wang, P. Dai, H. Fang, F. Wu, Q. Guo, P.K. Liaw, N. Hua, Novel Ti-Zr-Hf-Nb-Fe refractory high-entropy alloys for potential biomedical applications, J. Alloys Compd. 906 (2022) 164383, https://doi.org/10.1016/j.jallcom.2022.164383.

[7] L. Wu, G. Wei, G. Wang, H. Wang, J. Ren, Creating win-wins from strength–ductility trade-off in multi-principal element alloys by machine learning materials, Mater. Today Commun. 32 (2022) 104010, https://doi.org/10.1016/j.mtcomm.2022.104010.

[8] C. Chen, L. Ma, Y. Zhang, P.K. Liaw, J. Ren, Accelerating the design of high-entropy alloys with high hardness by machine learning based on particle swarm optimization, Intermetallics 154 (2023) 107819, https://doi.org/10.1016/j.intermet.2022.107819.

[9] F. Liu, P.K. Liaw, Y. Zhang, Recent Progress with BCC-structured high-entropy alloys, Metals 12 (3) (2022) 501, https://doi.org/10.3390/met12030501.

[10] Q. Xing, Y. Zhang, Amorphous phase formation rules in high-entropy alloys, Chinese Phys. B 26 (1) (2017) 018104, https://doi.org/10.1088/1674-1056/26/1/018104.

[11] K. Yao, L. Liu, J. Ren, Y. Guo, Y. Liu, Y. Cao, R. Feng, F. Wu, J. Qi, J. Luo, P.K. Liaw, W. Chen, High-entropy intermetallic compound with ultra-high strength and thermal stability, Scr. Mater. 194 (2021) 113674, https://doi.org/10.1016/j.scriptamat.2020.113674.

[12] C.L. Tracy, S. Park, D.R. Rittman, S.J. Zinkle, H. Bei, M. Lang, R.C. Ewing, W. L. Mao, High pressure synthesis of a hexagonal close-packed phase of the high-entropy alloy CrMnFeCoNi, Nat Commun. 8 (2017) 15634, https://doi.org/10.1038/ncomms15634.

[13] D. Beniwal, P.K. Ray, Learning phase selection and assemblages in high-entropy alloys through a stochastic ensemble-averaging model, Comput. Mater. Sci. 197 (2021) 110647, https://doi.org/10.1016/j.commatsci.2021.110647.

[14] J. Li, G. Zhou, J. Han, H. Zhang, Y. Peng, L. Chen, X. Cao, P.K. Liaw, Special hot working plastic deformation behavior and microstructure evolution mechanism of single-phase BCC structure AlFeCoNiMo0.2 highentropy alloy, J. Alloys Compd. 955 (2023) 170149, https://doi.org/10.1016/j.jallcom.2023.170149.

[15] R. Feng, P.K. Liaw, M.C. M. W, First-principles prediction of high-entropy-alloy stability. NPJ Comput. Mater. 3 (2017) 50. doi: 10.1038/s41524-017-0049-4.

[16] J. Yuan, Y. Liu, Z. Li, M. Wang, Q. Wang, C. Dong, Cluster-model-embedded first-principles study on structural stability of body-centered-cubic-based Ti-Zr-Hf-Nb Refractory high-entropy alloys, J. Phase Equilib. Diffus. 42 (2021) 647–655, https://doi.org/10.1007/s11669-021-00899-5.

[17] W.M. Choi, Y.H. Jo, S.S. Sohn, S. Lee, B.J. Lee, Understanding the physical metallurgy of the CoCrFeMnNi high-entropy alloy: an atomistic simulation study, NPJ Comput. Mater. 4 (1) (2018), https://doi.org/10.1038/s41524-017-0060-9.

[18] S. Chen, Z.H. Aitken, S. Pattamatta, Z. Wu, Z. Yu, D.J. Srolovitz, P.K. Liaw, Y. Zhang, Simultaneously enhancing the ultimate strength and ductility of high-entropy alloys via short-range ordering, Nat. Commun. 12 (1) (2021) 4953, https://doi.org/10.1038/s41467-021-25264-5.

[19] L. Fan, Y. Ji, G. Wang, J. Chen, K. Chen, X. Liu, Z. Wen, High entropy alloy electrocatalytic electrode toward alkaline glycerol valorization coupling with acidic hydrogen production, J. Am. Chem. Soc. 144 (16) (2022) 7224–7235, https://doi.org/10.1021/jacs.1c13740.

[20] M.C. Gao, D.E. Alman, Searching for next single-phase high-entropy alloy compositions, Entropy 15 (10) (2013) 4504–4519, https://doi.org/10.3390/e15104504.

[21] C. Varvenne, A. Luque, W.A. Curtin, Theory of strengthening in fcc high entropy alloys, Acta Mater. 118 (2016) 164–176, https://doi.org/10.1016/j.actamat.2016.07.040.

[22] A. Durga, K.C. Hari Kumar, B.S. Murty, Phase Formation in Equiatomic High Entropy Alloys: CALPHAD Approach and Experimental Studies, Trans. Indian. Inst Met. 65 (4) (2012) 375–380. doi: 10.1007/s12666-012-0138-5.

[23] O.N. Senkov, S.V. Senkova, C. Woodward, Effect of aluminum on the microstructure and properties of two refractory high-entropy alloys, Acta Mater. 68 (2014) 214–228, https://doi.org/10.1016/j.actamat.2014.01.029.

[24] Y. Li, W. Guo, Machine-learning model for predicting phase formations of high-entropy alloys, Phys. Rev. Mater. 3 (2019) 095005, https://doi.org/10.1103/PhysRevMaterials.3.095005.

[25] M.H. Tsai, R.C. Tsai, T. Chang, W. Huang, Intermetallic phases in high-entropy alloys: statistical analysis of their prevalence and structural inheritance, Metals 9 (2019) 247, https://doi.org/10.3390/met9020247.

[26] H. Chang, Y. Tao, P.K. Liaw, J. Ren, Phase prediction and effect of intrinsic residual strain on phase stability in high-entropy alloys with machine learning, J. Alloys Compd. 921 (2022) 166149, https://doi.org/10.1016/j.jallcom.2022.166149.

[27] L. Wang, P. Li, W. Zhang, F. Wan, J. Wu, L. Yong, X. Liu, Prediction of phase selection of amorphous alloys and high entropy alloys by artificial neural network, Comput. Mater. Sci. 223 (2023) 112129, https://doi.org/10.1016/j.commatsci.2023.112129.
[28] R. Jain, U. Lee, S. Samal, N. Park, Machine-learning-guided phase identification and hardness prediction of Al-Co-Cr-Fe-Mn-Nb-Ni-V containing high entropy alloys, J. Alloys Compd. 956 (2023) 170193, https://doi.org/10.1016/j.jallcom.2023.170193.
[29] W. Huang, P. Martin, H.L. Zhuang, Machine-learning phase prediction of high-entropy alloys, Acta Mater. 169 (2019) 225–236, https://doi.org/10.1016/j.actamat.2019.03.012.
[30] Y. Zhang, C. Wen, C. Wang, S. Antonov, D. Xue, Y. Bai, Y. Su, Phase prediction in high entropy alloys with a rational selection of materials descriptors and machine learning models, Acta Mater. 185 (2020) 528–539, https://doi.org/10.1016/j.actamat.2019.11.067.
[31] Y.V. Krishna, U.K. Jaiswal, R.M. R, Machine learning approach to predict new multiphase high entropy alloys, Scr. Mater. 197 (2021) 113804. doi: 10.1016/j.scriptamat.2021.113804.
[32] H. Ying, S. Liu, Z. Wu, W. Dong, J. Ge, H. Hahn, V. Provenzano, X. Wang, S. Lan, Phase selection rule of high-entropy metallic glasses with different short-to-medium-range orders, Rare Met. 41 (2022) 2021–2027, https://doi.org/10.1007/s12598-022-01973-8.
[33] W. Zhu, W. Huo, S. Wang, X. Wang, K. Ren, S. Tan, F. Fang, Z. Xie, J. Jiang, Phase formation prediction of high-entropy alloys: a deep learning study, J. Mater. Res. Technol. 18 (2022) 800–809, https://doi.org/10.1016/j.jmrt.2022.01.172.
[34] S.Y. Lee, S. Byeon, H.S. Kim, H. Jin, S. Lee, Deep learning-based phase prediction of high-entropy alloys: optimization, generation, and explanation, Mater. Des. 197 (2021) 109260, https://doi.org/10.1016/j.matdes.2020.109260.
[35] N. Qu, Y. Liu, Y. Zhang, D. Yang, T. Han, M. Liao, Z. Lai, J. Zhu, L. Zhang, Machine learning guided phase formation prediction of high entropy alloys, Mater. Today Commun. 32 (2022) 104146, https://doi.org/10.1016/j.mtcomm.2022.104146.
[36] S. Singh, N.K. Katiyar, S. Goel, S.N. Joshi, Phase prediction and experimental realisation of a new high entropy alloy using machine learning, Sci Rep 13 (1) (2023) 4811, https://doi.org/10.1038/s41598-023-31461-7.
[37] S. Hou, M. Sun, M. Bai, D. Lin, Y. Li, W. Liu, A hybrid prediction frame for HEAs based on empirical knowledge and machine learning, Acta Mater. 228 (2022) 117742, https://doi.org/10.1016/j.actamat.2022.117742.
[38] Y. Pan, L. Zhang, X. Wu, M.J. Skibniewski, Multi-classifiier information fusion in risk analysis, Inf. Fusion 60 (2020) 121–136, https://doi.org/10.1016/j.inffus.2020.02.003.
[39] V. Yaghoubi, L. Cheng, W. Van, Paepegem, M, Kersemans, CNN-DST: Ensemble Deep Learning Based on Dempster-Shafer Theory for Vibration-Based Fault Recognition, Struct Health Monit 21 (5) (2022) 2063–2082, https://doi.org/10.48550/arXiv.2110.07191.
[40] B. Xu, Y. Sun, Cutting-state identification of machine tools based on improved Dempster-Shafer evidence theory, Int. J. Adv. Manuf. Technol. 124 (2022) 4099–4106, https://doi.org/10.1007/s00170-022-09056-9.
[41] Y. Hua, Y. Sun, G. Xu, S. Sun, E. Wang, Y. Pang, A fault diagnostic method for oil-immersed transformer based on multiple probabilistic output algorithms and improved DS evidence theory, Int. J. Electr. Power Energy Syst. 137 (2022) 107828, https://doi.org/10.1016/j.ijepes.2021.107828.
[42] Y. Fan, Y. Zhang, B. Guo, X. Luo, Q. Peng, Z. Jin, A hybrid Sparrow Search algorithm of the Hyperparameter optimization in deep Learning, Mathematics 10 (16) (2022), https://doi.org/10.3390/math10163019.
[43] M. Mirza, S. Osindero, Conditional generative adversarial nets. arXiv, 2014, arXiv: 1411.1784. doi: 10.48550/arXiv.1411.1784.
[44] J. Xue, B. Shen, A novel swarm intelligence optimization approach: sparrow search algorithm, Syst. Sci. Control. Eng. 8 (1) (2020) 22–34, https://doi.org/10.1080/21642583.2019.1708830.
[45] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann Math Stat. 38 (2) (1967) 325–339.
[46] G. Shafer, A mathematical theory of evidence, Technometrics 20 (1) (1978) 106.
[47] Z. Liu, Q. Pan, J. Dezert, J.-W. Han, Y. He, Classifier fusion with contextual reliability evaluation, IEEE Trans Cybern 48 (5) (2018) 1605–1618, https://doi.org/10.1109/TCYB.2017.2710205.
[48] W. Zhang, Y. Deng, Combining conflicting evidence using the DEMATEL method, Soft Comput. 23 (2019) 8207–8216, https://doi.org/10.1007/s00500-018-3455-8.
[49] H. Yu, X. Wang, B. Ren, T. Zeng, M. Lv, C. Wang, An efficient Bayesian inversion method for seepage parameters using a data-driven error model and an ensemble of surrogates considering the interactions between prediction performance indicators, J. Hydrol. 604 (2022) 127235, https://doi.org/10.1016/j.jhydrol.2021.127235.
[50] Z. Deng, J. Wang, A Novel Evidence Conflict Measurement for Multi-Sensor Data Fusion Based on the Evidence Distance and Evidence Angle, Sensors 20 (2) (2020). doi: 10.3390/s20020381.
[51] F. Xiao, Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy, Inf. Fusion 46 (2019) 23–32, https://doi.org/10.1016/j.inffus.2018.04.003.
[52] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Stat. 38 (1967) 325–339, https://doi.org/10.1007/978-3-540-44792-4_3.
[53] G. Shafer, A mathematical theory of evidence, Technometrics 20 (1978) 106, https://doi.org/10.1080/00401706.1978.10489628.
[54] M.G. Poletti, L. Battezzati, Electronic and thermodynamic criteria for the occurrence of high entropy alloys in metallic systems, Acta Mater. 75 (2014) 297–306, https://doi.org/10.1016/j.actamat.2014.04.033.
[55] Z. Zhou, Y. Zhou, Q. He, Z. Ding, F. Li, Y. Yang, Machine learning guided appraisal and exploration of phase design for high entropy alloys, NPJ Comput. Mater. 5 (1) (2019) 128, https://doi.org/10.1038/s41524-019-0265-1.
[56] S. Guo, C. Liu, Phase stability in high entropy alloys: formation of solid-solution phase or amorphous phase, Progr. Natural Sci. Mater. Int. 21 (6) (2011) 433–446, https://doi.org/10.1016/S1002-0071(12)60080-X.
[57] A.F. Andreoli, J. Oravaa, P.K. Liawb, H. Weber, M.F. de Oliveirac, K. Nielschd, I. Kaban, The elastic-strain energy criterion of phase formation for complex concentrated alloys, Materialia 5 (2019) 100222, https://doi.org/10.1016/j.mtla.2019.100222.
[58] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, Improved techniques for training gans, in: Advances in Neural Information Processing Systems, 2016, p. 29, https://doi.org/10.48550/arXiv.1606.03498.
[59] S. Kullback, R.A. Leibler, On information and sufficiency, Ann Math Statistics. 22 (1) (1951) 79–86.
[60] S. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 4768–4777, https://doi.org/10.48550/arXiv.1705.07874.
[61] R. Tsaih, Sensitivity analysis, neural networks, and the finance, IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339), 1999, pp. 3830-3835 vol.6, https://doi: 10.1109/IJCNN.1999.830765.
[62] L. Sang, Y. Xu, Amorphous behavior of ZrxFeNiSi$_{0.4}$B$_{0.6}$ high entropy alloys synthesized by mechanical alloying, J. Non. Cryst. Solids. 530 (15) 2020 119854. doi: 10.1016/j.jnoncrysol.2019.119854.
[63] A. Inoue, Stabilization of metallic supercooled liquid and bulk amorphous alloys, Acta Mater. 48 (1) (2000) 279–306, https://doi.org/10.1016/S1359-6454(99)00300-6.
[64] X. Yang, Y. Zhang, Prediction of high-entropy stabilized solid-solution in multi-component alloys, Mater. Chem. Phys. 132 (2012) 233–238, https://doi.org/10.1016/j.matchemphys.2011.11.021.
[65] R. Chen, G. Qin, H. Zheng, L. Wang, Y. Su, Y. Chiu, H. Ding, J. Guo, H. Fu, Composition design of high entropy alloys using the valence electron concentration to balance strength and ductility, Acta Mater. 144 (2018) 129–137, https://doi.org/10.1016/j.actamat.2017.10.058.
[66] S. Guo, Q. Hu, C. Ng, C.T. Liu, More than entropy in high-entropy alloys: forming solid solutions or amorphous phase, Intermetallics 41 (2013) 96–103, https://doi.org/10.1016/j.intermet.2013.05.002.
[67] U. Subedi, Y.A. Coutinho, P.B. Malla, K. Gyanwali, A. Kunwar, Automatic featurization aided data-driven method for estimating the presence of Intermetallic phase in multi-principal element alloys, Metals 12 (2022) 964, https://doi.org/10.3390/met12060964.