Available online at www.sciencedirect.com
ScienceDirect
Trans. Nonferrous Met. Soc. China 35(2025) 3813−3823
Transactions of Nonferrous Metals Society of China
www.tnmsc.cn

Machine learning with center−environment attention mechanism for multi-component Nb alloys

Yu-chao TANG[1], Bin XIAO[2], Jian-hui CHEN[2], Shui-zhou CHEN[3], Yi-hang LI[2], Fu LIU[1], Wan DU[2], Yi-heng SHEN[2], Xue FAN[2], Quan QIAN[3], Yi LIU[1,2]

1. Shanghai Engineering Research Center for Integrated Circuits and Advanced Display Materials, College of Sciences, Shanghai University, Shanghai 200444, China;
2. Materials Genome Institute, Shanghai University, Shanghai 200444, China;
3. School of Computer Engineering & Science, Shanghai University, Shanghai 200444, China

Received 2 March 2024; accepted 10 September 2024

Abstract: The graph-based representation of material structures, along with deep neural network models, often lacks locality and requires large datasets, which are seldom available in specialized materials research. To address this challenge, we developed a more data-efficient center−environment (CE) structure representation that incorporates a predefined attention-focused mechanism. This approach was applied in a machine learning (ML) study to examine the local alloying effects on the structural stability of Nb alloys. In the CE feature model, the atomic environment type (AET) method was utilized, which effectively describes the low-symmetry physical shell structures of neighboring atoms. The optimized ML-CE_{AET} models successfully predicted double-site substitution energies in Nb with a mean absolute error of 55.37 meV and identified Si−M pairs (where M = Ta, W, Re, and lanthanide rare-earth elements) as promising stabilizers for Nb. The ML-CE_{AET} model's good transferability was further confirmed through accurate prediction of untrained alloying element Nb. Significantly, in cases involving small datasets, non-deep learning models with CE features outperformed deep learning models based on graph features reported in the literature.
Key words: machine learning; center−environment feature; atomic environment type; Nb alloy design

1 Introduction

The next generation aviation engines with higher push-weight ratio and efficiency require the ultra-high temperature structural materials beyond the current commercialized Ni-based superalloys. Many intermetallic compounds including NbSi-based superalloys appear as potential turbine materials owing to their high melting point (2400 °C) and relatively low density (6.6−7.2 g/cm^3) [1,2]. The NbSi binary phase diagram has a wide two-phase region of Nb and Nb_5Si_3 with eutectic and eutectoid reactions, offering various in situ composites microstructures [3,4].

Experimental and theoretical investigations show that alloying can effectively improve the room-temperature plasticity, high-temperature strength, and oxidation resistance of NbSi-based alloys [5]. The reported alloying elements in NbSi-based alloys mainly include Ti, Cr [6], Al, Hf [7], Zr, Sn, Mo, W, V, Ta, Fe, Zr, and B [8]. The NbSi-based superalloys have been gradually developed from the original NbSi binary system to the ternary and the multi-component systems [9]. BEWLAY et al [10] designed and fabricated the Nb−Ti−Hf−Cr−Al−Si alloys with excellent comprehensive properties after adding Ti and Hf

Corresponding author: Yi LIU, E-mail: yiliu@shu.edu.cn
https://doi.org/10.1016/S1003-6326(25)66914-7

elements; ZHANG and GUO [8] developed Nb−Ti−Si−Cr−Al−Hf−Zr and Nb−Ti−Si−Cr−Al−Hf−B superalloys. The high-throughput synthesis and oxidation experiments were used to explore the oxidation of Nb−M alloys with various combinations of component elements, concentration, oxidation temperature and time [5]. The trial-and-error experiments are costly and slow while the fundamental correlational alloying effects are still poorly understood for multi-component alloys [11,12].

The first-principles computations based on density functional theory (DFT) have been used to examine various properties and the micro-scopic strengthening mechanisms of NbSi-based superalloys [13]. The data-driven machine learning (ML) predictions can help to accelerate the expensive first-principles calculations [14−16]. We have recently developed center-environment (CE) feature models incorporating both compositional and structural information into ML features. The CE feature models introduce the elementary physiochemical features of environment atoms surrounding the center atoms, accounting for the influences of environments on the center atoms. LI et al [17,18] combined several single CE (sCE) atom sets into multiple CE (mCE) atom sets containing multiple center atoms and their associated environment atoms, which accurately predict the formation energies, lattice parameters, and band gaps of spinel and perovskite oxides. WANG et al [19] developed a surface center−environment (SCE) feature model and effectively predicted the adsorption free energies of intermediate species (HO*, O*, and HOO*) and the overpotentials of oxygen evolution reaction (OER) on the surfaces of perovskite oxides. CHEN et al [20] used the SCE feature model to predict the adsorption energy of C_2H_2 on various alloy surfaces to reveal the initial growth mechanism of nanocarbon materials. GUO et al [21,22] applied CE models with the nearest neighbor (NN) environment atoms to predict the formation energies and lattice constants of $L1_2$ Co_3(Al,X) with high symmetry. Despite the success of previous CE feature models, the NN definition for the environment atoms becomes tricky when applied to low symmetry crystals since the cutoff distance varies depending on local distorted structures. The incorporation of more than enough NN environment atoms does not necessarily improve the prediction accuracy. Instead, too large distance cutoff may introduce redundant negative effects since the CE is intrinsically a localized representation and too large cutoff may interfere with the other local CE atom sets. Therefore, the general definition of appropriate environment atoms is required in the CE feature construction especially for the complex low-symmetry crystal structures, becoming the major motivation of methodology development in this work. The broader impact of this work would be the alternative to the current graph-based neural network methods with complex architectures that require a large amount of expensive training data, limiting their practical applications in materials science [23,24].

The first-principles calculations were used to investigate systematically the alloying effects on the stability and mechanical properties of Nb alloys [25]. A large number of calculations from first-principles are still too costly to study many other alloying elements. Aiming to accelerate the studies of new alloying elements, the ML methods based on the previous first-principles computational data were developed to investigate the structural stability of the alloyed Nb phases in this work. Firstly, we developed the generalized CE feature model, specifically adapted for low-symmetry crystals, by examining different definitions of environment atoms and weights in feature constructions as well as ML algorithms. The optimized ML models were then applied without modification to predict the effects of untrained alloying elements, which were further validated by the additional DFT calculations.

2 Models and methods

2.1 Training dataset

The training datasets are built based on the first-principles calculations on the alloyed Nb [25]. Figure S1 of Supporting Materials (SM) depicts the experimental structures of Nb (body-centered cubic, BCC) crystals with the lattice parameters taken from the Materials Platform for Data Science (MPDS) [26]. The configurations of the studied substitution pair sites were depicted in Fig. 1 and Fig. S1 of SM. The numbers of substitution systems with non-equivalent pairs (Fig. 1) were listed in Table S1 of SM. Figure S2 of SM shows statistical

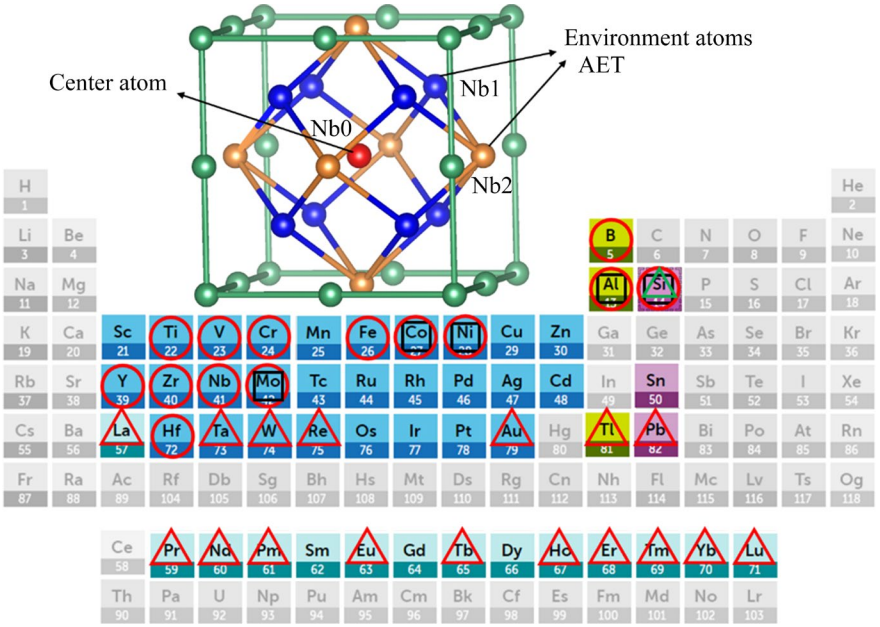

Fig. 1 Center−environment substitution site model with center atom (Nb0) and environment atoms (the first-nearest neighbor Nb1 and the second-nearest neighbor Nb2) defined by AET in Nb BCC convention cell (The studied substitution alloying elements in Nb are described in the periodic table: the red circles represent the 14 alloying elements in the DFT training dataset among which the black squares indicate the single-site stabilizers; the elements with the colored background are the 34 new elements studied by the ML models where the triangles represent the double-site stabilizers at X_{Nb0} (in green) and $X_{Nb1,2}$ (in red) sites in the double-site substitution systems $X_{Nb0}Y_{Nb1,2}$@Nb; the elements with the light grey backgrounds were not studied in this work)

Gaussian distributions of the target properties in Nb. Figure 1 and Fig. S3 of SM indicate the 14 substitution alloying elements in the periodic table.

Considering the single-site and double-site substitutions at the non-equivalent site pairs with the 14 alloying elements (Fig. 1 and Table S1 of SM), we collected 210 double-site substitution energies (E_{DS}) in Nb phase from the literature [25]. We also calculated the incremental single-site substitution energy (E_{SS}) in the cases of double-site substitution and the local bond length change (Δd) as defined in Text S1 of SM.

The CE feature models were constructed incorporating local structure and composition information (see more details in Table S2 of SM). The chemical composition (CC) models were also introduced for comparison (Text S2 of SM).

2.2 CE feature model

The CE features encoding the local structure and composition information have been applied to studying alloys, oxides, and surface catalyst reactions [17−22]. The CE composition−structure feature models were proposed in this work for complex substitution configurations with low symmetry. The CE feature model can be described as an n+1 dimensional compound features as follows:

$$D=[D_1, \cdots, D_i, \cdots, D_n, T] \ (n=20) \quad (1)$$

D consists of n elementary features of element or pure substance (D_i) and the target property T. D_i is a two-dimensional vector of the ith elementary property including the center and environment components defined as follows:

$$D_i=[d_{C,i}, d_{E,i}], \ i=1, 2, \cdots, n \quad (2)$$

$$d_{C,i}=p_{C,i} \quad (3)$$

$$d_{E,i}=\sum_{j=1}^{N}\left(w_{E,j}p_{E,j,i}\right) \quad (4)$$

$$w_{E,j}=r_j^m \Big/ \sum_{j=1}^{N} r_j^m \ (m=-1, 1/2) \quad (5)$$

where subscripts C and E represent the center atoms and environment atom, respectively; i is the elementary property index and j is the index of environment atoms; $p_{C,i}$ is the ith elementary property of the center atom; $p_{E,j,i}$ is the ith property

of the *j*th environment atom around the center atom; $w_{E,j}$ denotes the weight of elementary properties, expressed as a function of normalized distance r_j between the center atom and the *j*th environment atom. The weight is inversely proportional to the distance as r_j^m (m=−1, −1/2) where different powers m were studied and compared in this work.

It is well known that feature engineering determines the accuracy of ML modeling [17,27−30]. The CE features are compound features consisting of an assembly of elementary property features encoded with the local structural information specified by the center and environment atoms: (1) Elementary property features are various elementary physicochemical properties readily available from the fundamental database [31], e.g., atomic mass, radius, electronegativity, and the number of valence electrons of elements as well as density, melting temperature, and bulk modulus of pure substance among others. In total 40 elementary properties were adopted in the feature construction as listed in Table S2 of SM. (2) Compound property features are constructed by a linear combination of the elementary properties of the center atom or the environment atoms with weights inversely proportional to the distance between the center atom and the environment atom (r_j^m, m=−1, −1/2). By this way, CE features can encode the elementary properties with the local composition and structure information, providing a general digital representation of materials structure. For comparison with the CE feature models, the CC feature models were constructed by considering chemical composition only without structure information. The CC feature construction is similar to that of CE except that the weight is independent of distance r_j^m (m=0).

2.3 Machine learning algorithms and evaluation

The machine learning algorithms adopted the support vector regression (SVR) algorithm [32] with a radial basis function (rbf) kernel function and random forest (RF) [33] implemented in the Scikit-learn library of Python. The hyper-parameters of SVR and RF algorithms were determined by grid search methods as described in Text S3 and Table S3 of SM.

The training datasets and the test datasets were constructed by random 8∶2 split of the original dataset 20 times. In each split, the 80% training dataset was subjected to 5-fold cross-validation of the training model. The original 20% test dataset was used independently to evaluate the trained ML model. To evaluate the performance of the regression models, the statistical metrics adopted correlation coefficient (R^2), the mean absolute error (MAE), and the root means square error (RMSE) as defined in Eqs. (6)−(8), respectively. To obtain statistical results, the performance results were averaged over the 20 ML models.

$$R^2 = 1 - \frac{\sum_{j=0}^{n-1}(\hat{y}_j - y_j)^2}{\sum_{j=0}^{n-1}(\bar{y}_j - y_j)^2} \tag{6}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right| \tag{7}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \tag{8}$$

where n is the number of samples; y_j is the true value; $\hat{y}_j$ is the predicted value; $\bar{y}_j$ is the mean of predicted values.

3 Results and discussion

3.1 Construction of machine learning models

3.1.1 CE_{NN} and CE_{AET} feature models

The CE feature model essentially provides a center and environment framework of encoding local composition and structure information of materials. The center atoms are normally the focused critical atoms, e.g., the substitution alloying elements at the non-equivalent sites of Nb in this work. It is physically necessary to consider the effects of environment atoms on the center atoms. The definition of environment atoms is critical to the appropriate representation of local chemical and structural information. To explore the impact of the environment atoms on the performance of ML-CE models, we developed two construction methods of environment atoms described as follows.

(1) Nearest neighbor (dubbed CE_{NN}) feature model. For high symmetry crystal materials, it is natural to select the environment atoms based on the distances between the center and environment atoms. In the nearest neighbor CE_{NN} feature models, the environment atoms are defined to include the

*n*th-nearest neighbor atoms from the center atom. In this work, up to the fifth nearest neighbor atoms were considered as the environment atoms in the alloyed Nb.

(2) AET (dubbed CE_{AET}) feature model. For the crystal structures with low symmetry or chemical orderings in complex multi-component high-entropy materials, the distance-based cutoff definition is no longer appropriate to describe the environment. In this work, we adopted a physics-based environment atom definition for CE feature construction using the AET concept originally proposed by DAAMS et al [34] for crystal classification. The AET represents a complete closed physical shell around a center atom based on the geometrical topology rather than distance cutoff only.

In the CE_{AET} representation, the AET environment atoms need to satisfy the maximum distance distribution gap and the convex volume rules. The rule of maximum distance distribution gap requires that the AET atoms have the maximum gap from the farther atoms in the nearest-neighbor histogram (NNH) of distance distributions, and a plot of the number of samples (n) versus the normalized distances between the center and the surrounding atoms (d/d_{min}, d is a given interatomic distance, and d_{min} is the shortest interatomic distance) is shown in Fig. 2. The second convex volume rule requires that the AET atoms must encompass a convex polyhedron.

Figure 2 depicts the AET cluster models and their NNHs with the centers of non-equivalent sites in Nb. The NNH of Nb is plotted in Fig. 2 where the AET cluster model is a rhombic dodecahedron cluster with its coordination number (CN=14) and polyhedron code ($8^{0.3}6^{0.4}$). The polyhedron code ($8^{0.3}6^{0.4}$) in Fig. 2 represents eight vertices adjoining no triangles and three squares as well as six vertices adjoining no triangles and four squares. The numbers of AET atoms vary depending on the local symmetry, so it is hard to predefine the *n*th nearest neighbors without individual check. The inappropriate choice of the *n*th nearest neighbors as the environment atoms may lead to an incomplete or redundant shell atoms and physically less meaningful features in the CE feature construction.

3.1.2 Prediction performance of ML models

To compare the prediction accuracy of different ML models, we show the performance metrics of the CE_{NN}, $CE_{AET,}$ and CC feature models with different weights and algorithms for Nb in Tables S4−S6 and Figs. S4−S7 of SM.

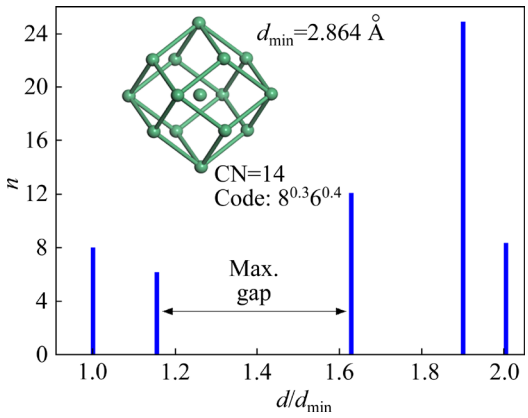

Fig. 2 Nearest-neighbor histogram (NNH) and atom environment type (AET) cluster model of Nb (The coordination polyhedron is a rhombic dodecahedron)

The SVR algorithm exhibited generally more accurate prediction by 100−200 meV than the RF algorithm with all studied features so that the SVR results were mainly used for discussion. The CE feature models (Tables S4 and S5 of SM) performed much better than the composition CC models (Table S6 of SM), indicating that the inclusion of structural information into the feature construction via CE framework is critical to describing the complex crystal structures in ML prediction. In addition, we also compared CE predictions with the other state-of-art deep learning ML models using graph theory features in the literatures [23,24,35]. The results show that CE models performed significantly better than the deep learning models [23,24,35] in the case of small dataset (Table 1).

Furthermore, the CE_{AET} models using the AET environment atoms had better prediction accuracy than the CE_{NN} models using the NN atoms even though more atoms may be included in the latter cases. This suggests that the physically closed shell is more appropriate to define ML features than the distance-based cutoff selection possibly with either insufficient or redundant environment atoms. From the comparison among the CE_{AET} feature models (Fig. 3), the weight at r_j^{-1} performed mostly better than that at $r_j^{-1/2}$, indicating that the linear combination of elementary property features with the weight of reciprocal distance is a reasonable choice

Table 1 Prediction performances of substitution energies of Nb alloys using various CC and CE features models and other deep machine learning models in literature

Model	R^2	MAE/meV	RMSE/meV
GCN [35]	0.90	163.80	252.20
GAT [23]	0.75	295.10	390.10
ALIGNN [24]	0.03	644.10	777.40
CC-RF	0.84	171.31	271.83
CE_{NN}-RF	0.86	182.38	245.08
CE_{AET}-RF	0.91	199.64	248.11
CC-SVR	0.94	146.80	176.92
CE_{NN}-SVR	0.98	79.87	98.61
CE_{AET}-SVR	0.99	55.37	91.85

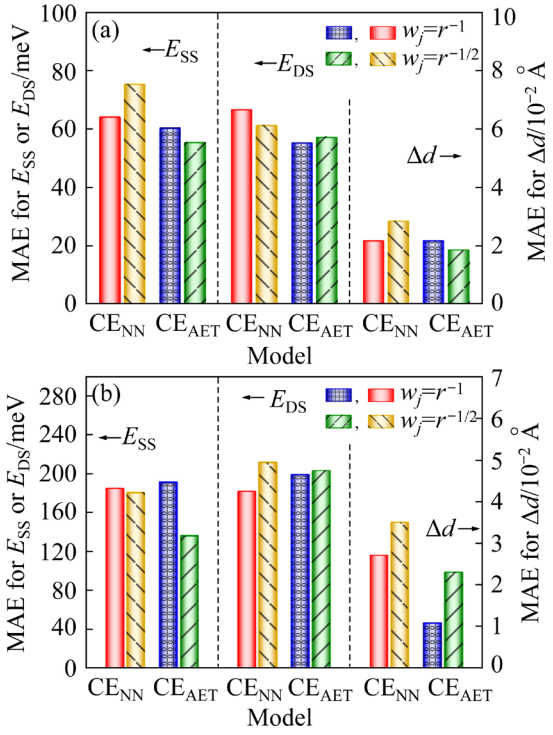

Fig. 3 MAE of prediction of Nb by SVR (a) and RF (b) methods using CE_{NN} and CE_{AET} feature models with different weights of r_j^m (m=−1, −1/2)

probably due to the r^{-1} distance scaling law of long range electrostatic interactions. Based on the comparison above, the optimal CE_{AET}-SVR models with weight $w_j = r^{-1}$ were mainly used to predict the target properties (E_{SS}, E_{DS}, and Δd) of untrained elements and structures of Nb alloys hereafter.

In summary, the comparison between various CE construction methods showed that: (1) the AET definition of environment atoms (CE_{AET}) performed better than the NN ones (CE_{NN}); (2) the weight of reciprocal distance (r^{-1}) behaved better in the linear combination of elementary features; (3) the SVR algorithm was slightly better than RF for substitution energy prediction.

The substitution energies and geometries of Nb predicted by the optimal CE_{AET}-SVR model (w_j=r^{-1}) are compared with the DFT results (Fig. 4). The ML prediction with other different features and algorithms are shown in Figs. S4−S6 of SM. The performance of E_{SS} in the train/test datasets predicted by the CE_{AET}-SVR model (Fig. 4) is R^2=0.99/0.99, MAE=27.87 meV/58.53 meV, and RMSE=68.01 meV/93.89 meV; the corresponding prediction performance of E_{DS} is R^2=0.99/0.99, MAE=34.84 meV/55.37 meV, and RMSE= 77.81 meV/91.85 meV; The prediction performance of Δd is R^2=0.95/0.82, MAE=0.71×10^{-2} Å/ 1.88×10^{-2} Å, RMSE=2.63×10^{-2} Å/4.57×10^{-2} Å, respectively. The MAE of double-site substitution energies predicted by the CE_{AET}-SVR models with different weights was ~50 meV, which was 10−25 meV smaller than that of the CE_{NN} models (Tables S4 and S5 of SM). The prediction accuracy of SVR models was much better than that of the RF algorithm (~200 meV). For the Δd, the CE_{AET} model prediction with the SVR algorithm slightly outperformed that with the RF algorithm. The studied substitution elements in Nb mostly caused slight expansion where most of Δd values were <0.3×10^{-2} Å.

To understand the site dependence of substitution energies, we plot the heat maps of the double-site substitution energy E_{DS} projection on the substitution pair sites (X_{Nb0}, Y_{Nb1}) and (X_{Nb0}, Y_{Nb2}) in Fig. 5 using the DFT and ML results, respectively, where the alloying elements are sorted by the metal radii. The distribution patterns of substitution energy predicted by the ML appear very similar to those by the DFT, confirming the reliable accuracy of the ML predictions. Such site-energy heat maps help to find the stabilizing element pairs quickly. For example, Si and Al substitutions stabilize Nb while Hf decreases the stability of Nb, confirmed by both DFT and ML. In summary, the machine learning method was validated against DFT and will be used to find new favorable stabilizing alloying elements in Nb alloys next.

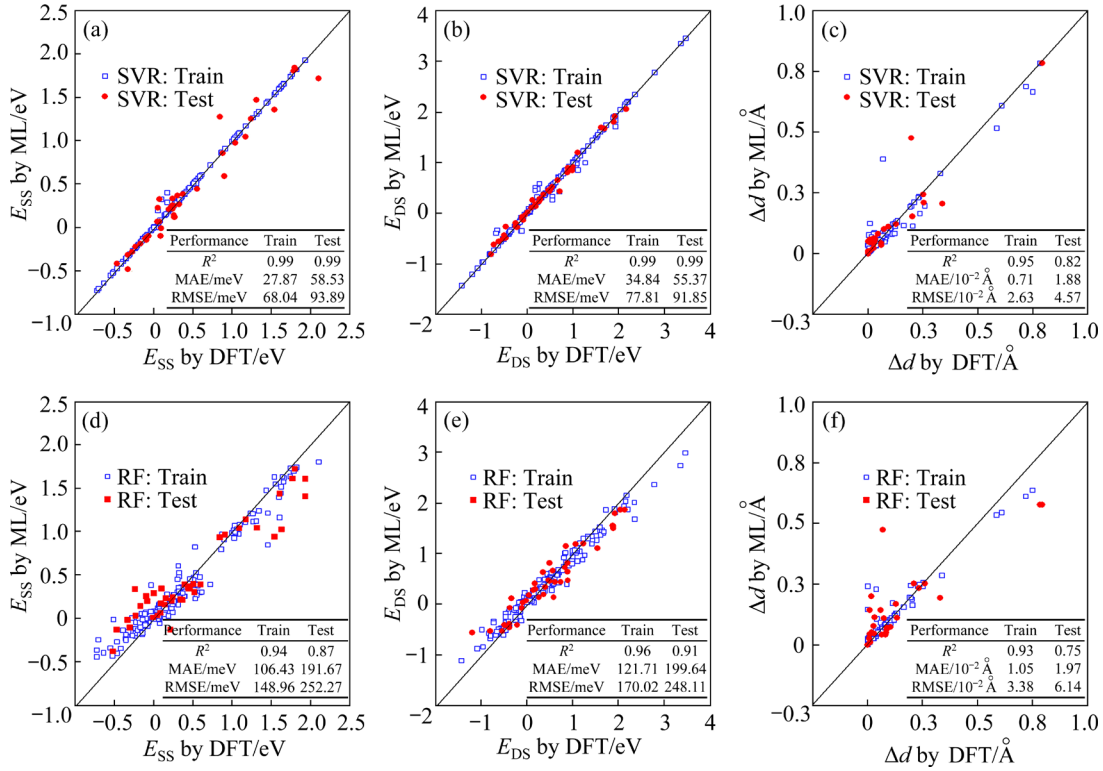

Fig. 4 E_{SS} (a, d), E_{DS} (b, e), and Δd (c, f) of Nb predicted by CE_{AET} models with SVR and RF algorithms with weight $w_j=r^{-1}$

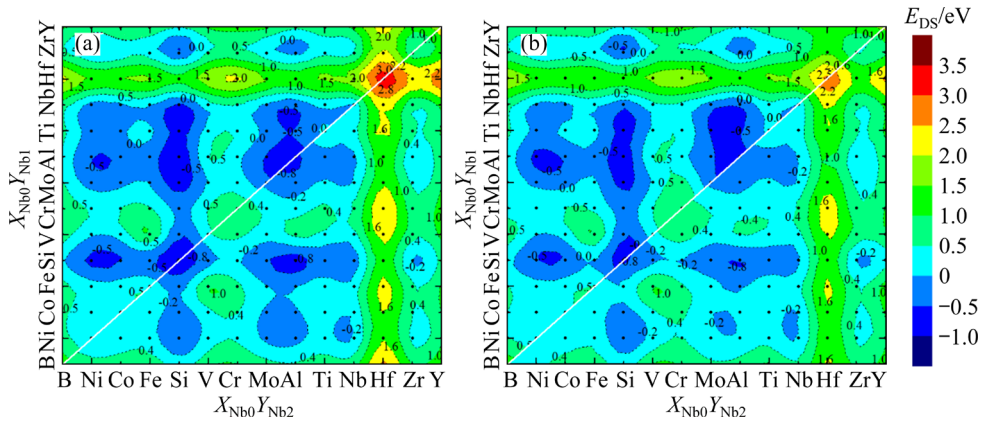

Fig. 5 Heat maps of double-site substitution energies E_{DS} of Nb predicted by DFT (a) and ML (b) projected on substitution site pairs (X_{Nb0}, Y_{Nb1}) and (X_{Nb0}, Y_{Nb2}) (X, Y= B, Ni, Co, Fe, Si, V, Mo, Al, Ti, Nb, Hf, Zr, and Y, sorted by increasing order of metal radii; The numbers 0, 1, and 2 denote the central Nb site, and the first- and second nearest-neighbor Nb sites of the environment atoms, respectively (shown in Fig. 1))

3.2 Applications of machine learning models

After the construction and comparison of the various ML models discussed above, the optimal CE_{AET}-SVR models with weight $w_j=r^{-1}$ were applied to predicting the unknown systems with new alloying elements. The transferability of ML predictions would significantly extend the prediction capability and efficiency beyond expensive first-principles computations.

3.2.1 Leave-p-out prediction of new alloying elements

To examine the capability of the ML models to predict the energy and structure of the new alloying elements, we predicted the E_{SS}, E_{DS}, and Δd for each of the 14 substituted alloying elements in the Nb phases using the leave-p-out cross validation method. Specifically, the full datasets containing the 14 elements were split into the test datasets of a target element and the training datasets of the

remaining 13 elements. The ML model trained with the 13-element dataset was used to predict the properties of the 14th element. Such leave-p-out validation procedures were performed for each of the 14 substitution elements. The R^2 and MAE of the leave-p-out ML prediction for the 14 alloying elements in Nb are shown in Fig. 6(a) and Fig. S8 of SM.

Figure 6(a) summarizes the MAE of the substitution energies of Nb in the leave-p-out prediction of each of the 14 alloying elements using CE_{AET}-SVR models. The MAE values of substitution energies for Ti, Zr, V, Nb, Fe, Co, and Ni elements were less than 300 meV; those for B, Al, Cr, and Mo elements were 300−600 meV; those for Si, Y, and Hf were larger than 600 meV. Figure S8 of SM shows the performance metrics of E_{DS} in Nb phase predicted by the CE_{AET}-SVR models. The

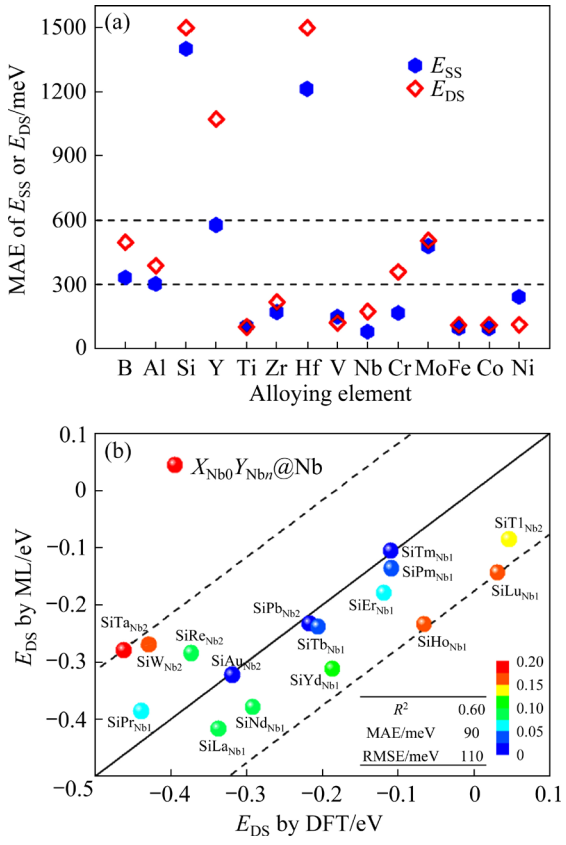

Fig. 6 (a) MAE of substitution energies E_{SS} in leave-p-out prediction of each of 14 alloying elements in Nb phase using CE_{AET}-SVR models (The alloying elements are sorted by the number of valence electrons); (b) Double-site substitution energies E_{DS} of stable double-site substitution systems $X_{Nb0}Y_{Nb1,2}$@Nb predicted by ML models (CE_{AET}-SVR) and DFT (The colored scale bar indicates the absolute errors from low (in blue) to high (in red))

R^2 of Co, Fe, Ni, Nb, Ti, V, and Zr reached 0.97, 0.98, 0.97, 0.96, 0.98, 0.94, and 0.94, respectively. The corresponding MAE values were 112.49, 112.73, 114.31, 175.39, 102.78, 187.75, and 220.12 meV, respectively. The other elements had larger MAE with R^2 less than 0.85.

The prediction errors of the new elements should be kept in mind in the application of the ML models. On the other hand, the prediction uncertainty of new elements may be physically meaningful since the deviation can reflect the difference or similarity of alloying effects between the predicted unknown and existing known elements. Indeed, the alloying elements with large prediction errors were the non-metallic main group elements Si as well as Y and Hf with a large metal radius, differing obviously from the other transition metal elements.

3.2.2 Prediction of new elements beyond training datasets

To evaluate the prediction capability of new elements beyond the DFT training datasets, we screened for the new stabilizing alloying elements in the double-site substituted Nb systems using the optimal CE_{AET}-SVR models. The 34 new candidate elements and their roles in stabilization are described in Fig. 1. In the periodic table, the red circles represent the 14 alloying elements in the DFT training dataset among which the black squares (Al, Si, Mo, Co, and Ni) indicate the single-site stabilizers; the elements with the colored background are the 34 new elements studied by the ML models where the triangles represent the stabilizers at X_{Nb0} (Si in green) and $X_{Nb1,2}$ (La, Hf, Ta, W, Re, Au, Tl, Pb, and rare earth elements in red) sites in the double-site substitution systems $X_{Nb0}Y_{Nb1,2}$@Nb.

The double-site substitution energies (E_{DS}) of 2046 double-site substitution systems $X_{Nb0}Y_{Nb1,2}$@Nb containing 14 trained and 34 untrained alloying elements were predicted by the optimal CE_{AET}-SVR models and plotted as functions of metallic radii in Fig. S9 of SM.

The optimal ML models predicted that the stabilized substitution configurations were $X_{Nb0}Y_{Nbn}$ (X_{Nb0}=Si; Y_{Nb1}=Pr, La, Nd, Tb, Lu, Yb, Er, Pm, Tm, and Ho; Y_{Nb2}=Ta, W, Re, Au, Pb, and Tl). The co-doping stabilizing pairs are Si−M (M=Ta, W, Re, Au, Pb, Tl, Pr, La, Nd, Tb, Lu, Yb, Er, Pm, Tm, and Ho). We found that all these stabilized systems

contain the leading Si element, consistent with our previous DFT predictions [25]. To validate these ML predictions, the E_{DS} values of these stable systems were calculated using DFT as shown in Fig. 6(b). The comparison shows the reasonable agreement between the ML and DFT results except for the $Si_{Nb}Ta_{Nb2}$, $Si_{Nb}Ho_{Nb1}$, and $Si_{Nb}Lu_{Nb1}$ pairs with large deviations. These extended ML predictions were validated further by the first-principles calculations, demonstrating the reliable transferability of ML prediction using the CE feature models. The efficient large-sale ML screening followed by the DFT validation on the limited promising candidates serves as an efficient acceleration design strategy.

4 Conclusions

(1) To conduct a machine learning study on local alloying effects, CE_{AET} models were developed as a general feature model for complex crystal structures. These models demonstrated effectiveness, efficiency, and transferability in predicting the alloying effects on the structural stability of Nb alloys.

(2) The AET definition of environment atoms (CE_{AET}) outperformed the nearest neighbor approach (CE_{NN}) when comparing various CE construction methods. In particular, the use of reciprocal distance weighting proved more effective in the linear combination of elementary features. Additionally, the ML-CE_{AET} models outperformed deep learning models that utilized graph-based features.

(3) The optimal CE_{AET}-SVR models predicted double-site substitution energies in Nb with a mean absolute error (MAE) of approximately 50 meV. These models were subsequently applied to predicting the substitution energies of untrained alloying elements in Nb, identifying new stabilizing pairs, Si−M (M=Ta, W, Re, Au, Pb, Tl, Pr, La, Nd, Tb, Lu, Yb, Er, Pm, Tm, and Ho). These predictions were further validated by the first-principles calculations, demonstrating the reliable transferability of ML predictions using CE feature models for the composition design of multi-component alloys.

CRediT authorship contribution statement

Yu-chao TANG: Methodology, Software, Investigation, Data curation, ML modelling, Visualization, Writing − Original draft; **Bin XIAO:** ML modelling, Software, Discussion; **Jian-hui CHEN:** ML modelling, Discussion; **Shui-zhou CHEN:** Software, Discussion; **Yi-hang LI:** ML modelling, Software; **Fu LIU:** Discussion, Data analysis; **Wan DU:** Discussion; **Yi-heng SHEN:** Software, Discussion; **Xue FAN:** Discussion; **Quan QIAN:** Discussion; **Yi LIU:** Conceptualization, Supervision, Methodology, Writing − Review & editing.

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.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 52373227, 52201016), the National Key Research and Development Program of China (Nos. 2017YFB0702901, 2017YFB0701502, 2023YFB4606200), Shanghai Technical Service Center for Advanced Ceramics Structure Design and Precision Manufacturing, China (No. 20DZ2294000), and Key Program of Science and Technology of Yunnan Province, China (No. 202302AB080020).

Supplementary Materials

Supplementary Materials in this paper can be found at: http://tnmsc.csu.edu.cn/download/17-p3813-2024-0348-Supplementary_Materials.pdf.

References

[1] POLLOCK T M. Alloy design for aircraft engines [J]. Nature Materials, 2016, 15(8): 809−815.

[2] LIU Wei, HUANG Shuai, YE Cheng-tong, JIA Li-na, KANG Yong-wang, SHA Jiang-bo, CHEN Bing-qing, WU Yu, XIONG Hua-ping. Progress in Nb−Si ultra-high temperature structural materials: A review [J]. Journal of Materials Science & Technology, 2023, 149: 127−153.

[3] HE Jia-hua, GUO Xi-ping, QIAO Yan-qiang. Oxidation and hot corrosion behaviors of Nb−Si based ultrahigh temperature alloys at 900 °C [J]. Transactions of Nonferrous Metals Society of China, 2021, 31: 207−221.

[4] TSAKIROPOULOS P. Alloys for application at ultra-high temperatures: Nb-silicide in situ composites [J]. Progress in Materials Science, 2022, 123: 100714.

[5] SHU Jin-tao, DONG Zi-qiang, ZHENG Chen, SUN An-kang, YANG Shuang, HAN Tao, LIU Yan-jie, WANG Zi-han, WANG Shu-juan, LIU Yi. High-throughput experiment-assisted study of the alloying effects on oxidation of Nb-based alloys [J]. Corrosion Science, 2022, 204: 110383.

[6] ZHAO J C, JACKSON M R, PELUSO L A. Determination of the Nb−Cr−Si phase diagram using diffusion multiples [J]. Acta Materialia, 2003, 51(20): 6395−6405.

[7] WANG Qi, ZHAO Tian-zhi, CHEN Rui-run, WANG Xiao-wei, XU Qin, WANG Shu, FU Heng-zhi. Collaborative optimization of microstructure, phase composition and room-temperature fracture toughness of Nb−Si based alloys using Ti, Zr and Hf elements [J]. Transactions of Nonferrous Metals Society of China, 2024, 34(1): 194−202.

[8] ZHANG Song, GUO Xi-ping. Alloying effects on the microstructure and properties of Nb−Si based ultrahigh temperature alloys [J]. Intermetallics, 2016, 70: 33−44.

[9] BEWLAY B P, JACKSON M R, ZHAO J C, SUBRAMANIAN P R, MENDIRATTA M G, LEWANDOWSKI J J. Ultrahigh-temperature Nb-silicide-based composites[J]. MRS Bulletin, 2003, 28(9): 646−653.

[10] BEWLAY B P, JACKSON M R, LIPSITT H A. The balance of mechanical and environmental properties of a multielement niobium-niobium silicide-based in situ composite [J]. Metallurgical and Materials Transactions A, 1996, 27(12): 3801−3808.

[11] LIU Yi, WANG Jiong, XIAO Bin, SHU Jin-tao. Accelerated development of hard high-entropy alloys with data-driven high-throughput experiments [J]. Journal of Materials Informatics, 2022, 2: 3.

[12] WEN Zhu-hao, LIN Hao-qin, CHEN Wei-min, BAI Ke-wu, ZHANG Li-jun. High-throughput exploration of composition-dependent elasto-plastic and diffusion properties of refractory multi-element Ti−Nb−Zr−W alloys [J]. Transactions of Nonferrous Metals Society of China, 2023, 33(9): 2646−2659.

[13] SHI Song-xin, ZHU Ling-gang, ZHANG Hu, SUN Zhi-mei, AHUJA R. Mapping the relationship among composition, stacking fault energy and ductility in Nb alloys: A first-principles study [J]. Acta Materialia, 2018, 144: 853−861.

[14] JUAN Yong-fei, NIU Guo-shuai, YANG Yang, XU Zi-han, YANG Jian, TANG Wen-qi, JIANG Hai-tao, HAN Yan-feng, DAI Yong-bing, ZHANG Jiao, SUN Bao-de. Accelerated design of Al−Zn−Mg−Cu alloys via machine learning [J]. Transactions of Nonferrous Metals Society of China, 2024, 34(3): 709−723.

[15] MENON N, MONDAL S, BASAK A. Linking processing parameters with melt pool properties of multiple nickel-based superalloys via high-dimensional Gaussian process regression [J]. Journal of Materials Informatics, 2023, 3: 7.

[16] XI Sheng-kun, YU Jin-xin, BAO Long-ke, CHEN Liu-ping, LI Zhou, SHI Rong-pei, WANG Cui-ping, LIU Xing-jun. Machine learning-accelerated first-principles predictions of the stability and mechanical properties of L12-strengthened cobalt-based superalloys[J]. Journal of Materials Informatics, 2022, 2: 15.

[17] LI Yi-hang, XIAO Bin, TANG Yu-chao, LIU Fu, WANG Xiao-meng, YAN Fei-nan, LIU Yi. Center-environment feature model for machine learning study of spinel oxides based on first-principles computations [J]. The Journal of Physical Chemistry C, 2020, 124(52): 28458−28468.

[18] LI Yi-hang, ZHU Rui-jie, WANG Yuan-qing, FENG Ling-yan, LIU Yi. Center-environment deep transfer machine learning across crystal structures: From spinel oxides to perovskite oxides [J]. NPJ Computational Materials, 2023, 9(1): 109.

[19] WANG Xiao-meng, XIAO Bin, LI Yi-hang, TANG Yu-chao, LIU Fu, CHEN Jian-hui, LIU Yi. First-principles based machine learning study of oxygen evolution reactions of perovskite oxides using a surface center-environment feature model [J]. Applied Surface Science, 2020, 531: 147323.

[20] CHEN Rong, LIU Fu, TANG Yu-chao, LIU Yan-jie, DONG Zi-qiang, DENG Zhen-yan, ZHAO Xin-luo, LIU Yi. Combined first-principles and machine learning study of the initial growth of carbon nanomaterials on metal surfaces [J]. Applied Surface Science, 2022, 586: 152762.

[21] GUO Jing, XIAO Bin, LI Yi-hang, ZHAI Dong, TANG Yu-chao, DU Wan, LIU Yi. Machine learning aided first-principles studies of structure stability of $Co_3(Al,X)$ doped with transition metal elements [J]. Computational Materials Science, 2021, 200: 110787.

[22] GUO Jing, XIAO Bin, TANG Yu-chao, LI Yi-hang, ZHAI Dong, FAN Xue, LIU Yi. Element-configuration dependent first-principles machine learning studies of multiple alloying effects on the structure stability of $Co_3(Al,W)$ [J]. Computational Materials Science, 2024, 233: 112767.

[23] VELICKOVIC P, CUCURULL G, CASANOVA A, ROMERO A, LIO P, BENGIO Y. Graph attention networks [EB/OL]. [2018−06−19]. https://doi.org/10.48550/arXiv.1710.10903.

[24] CHOUDHARY K, DECOST B. Atomistic line graph neural network for improved materials property predictions [J]. NPJ Computational Materials, 2021, 7(1): 185.

[25] TANG Yu-chao, XIAO Bin, CHEN Jian-hui, LIU Fu, DU Wan, GUO Jing, LIU Yan-jie, LIU Yi. Multi-component alloying effects on the stability and mechanical properties of Nb and Nb−Si alloys: A first-principles study [J]. Metallurgical and Materials Transactions A, 2023, 54(2): 450−472.

[26] BLOCHL E, VILLARS P. Handbook of materials modeling [M]. Berlin: Springer, 2018.

[27] WARD L, LIU R Q, KRISHNA A, HEGDE V I, AGRAWAL A, CHOUDHARY A, WOLVERTON C. Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations [J]. Physical Review B, 2017, 96(2): 024104.

[28] ISAYEV O, OSES C, TOHER C, GOSSETT E, CURTAROLO S, TROPSHA A. Universal fragment descriptors for predicting properties of inorganic crystals [J]. Nature Communications, 2017, 8(1): 15679.

[29] OUYANG R H, CURTAROLO S, AHMETCIK E, SCHEFFLER M, GHIRINGHELLI L M. SISSO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates [J]. Physical Review Materials, 2018, 2(8): 083802.

[30] ZHANG Hui-ran, HU Rui, LIU Xi, LI Sheng-zhou, ZHANG Guang-jie, QIAN Quan, DING Guang-tai, DAI Dong-bo. An end-to-end machine learning framework exploring phase formation for high entropy alloys [J]. Transactions of Nonferrous Metals Society of China, 2023, 33(7): 2110−2120.

[31] STOLYARENKO A. Database on properties of chemical elements [EB/OL]. [2023−02−16]. https://phases.imet-db.ru/elements/main.aspx.
[32] DUCKER H, BURGES C, KAUFMAN L, SMOLA A, VAPNIK V. Support vector regression machines [J]. Advances in Neural Information Processing Systems, 1997, 28(7): 779−784.
[33] LEO B. Random forests [J]. Machine Learning, 2001, 45(1): 5−32.
[34] DAAMS J L C, van VUCHT J H N, VILLARS P. Atomic-environment classification of the cubic “Intermetallic” structure types [J]. Journal of Alloys and Compounds, 1992, 182(1): 1−33.
[35] LOUIS S Y, ZHAO Y, NASIRI A, WANG X R, SONG Y Q, LIU F, HU J J. Graph convolutional neural networks with global attention for improved materials property prediction [J]. Physical Chemistry Chemical Physics, 2020, 22(32): 18141−18148.

多组元铌合金的中心−环境注意力机制机器学习

唐宇超[1]，肖 斌[2]，陈建辉[2]，陈水洲[3]，李一航[2]，
刘 馥[1]，杜 婉[2]，沈祎恒[2]，范 雪[2]，钱 权[3]，刘 轶[1,2]

1. 上海大学 理学院 上海市集成电路和先进显示材料工程研究中心，上海 200444；
2. 上海大学 材料基因工程组研究院，上海 200444；
3. 上海大学 计算机工程与科学学院，上海 200444

摘 要：利用基于图的深度神经网络进行材料结构的数字编码方法缺少局部性特征，且需要大量数据，这在特定的材料研究中很难满足。本文提出了一种数据利用效率更高的中心−环境(CE)结构特征表示方法，利用预先定义的注意力集中机制构建机器学习模型，研究局部合金化对 Nb 合金结构稳定性的影响。本研究中的 CE 特征模型采用原子环境类型(AET)方法，很好地描述了相邻原子低对称性的物理壳层结构。优化的 ML-CE_{AET} 模型预测了 Nb 的双位点置换能，其平均绝对误差为 55.37 meV，表明 Si−M(M=Ta, W, Re 和镧系稀土元素)置换对是 Nb 的稳定剂。通过直接预测未训练的 Nb 合金元素，进一步验证了 ML-CE_{AET} 模型的良好可迁移性。在小数据集的情况下，具有 CE 特征的非深度学习模型比文献中具有基于图特征的深度学习模型表现更好。

关键词：机器学习；中心−环境特征；原子环境类型；Nb 合金设计

(Edited by Wei-ping CHEN)