Materials Today Communications 26 (2021) 101871

Contents lists available at ScienceDirect

Materials Today Communications

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

Yield strength prediction of high-entropy alloys using machine learning

Uttam Bhandari *, Md. Rumman Rafi, Congyan Zhang, Shizhong Yang *

Department of Computer Science, Southern University and A&M College, Baton Rouge, Louisiana, 70813, United States

ARTICLE INFO

Keywords:
High entropy alloys
Random forest model
Yield strength prediction
MoNbTaTiW
HfMoNbTaTiZr

ABSTRACT

Yield strength at high temperature is an important parameter in the design and application of high entropy alloys (HEAs). However, the experimental measurement of yield strength at high temperature is quite costly, complicated, and time-consuming. Therefore, it is essential to identify and apply a robust method for the accurate prediction of yield strength at high temperature from the available experimental and simulation data. In this study, for the first time, a machine learning (ML) method based on the regression technique of random forest (RF) regressor is used to predict the yield strength of HEAs at the desired temperature. The yield strengths of MoNbTaTiW and HfMoNbTaTiZr at 800 °C and 1200 °C, are predicted using the RF regressor model. We find that the results are consistent with the experimental reports, showing that the RF regressor model predicts the yield strength of HEAs at the desired temperatures with high accuracy.

1. Introduction

High entropy alloys (HEAs) have been drawing the interest of many researchers after the invention of HEA by Yeh et al. [1] in 2004. Unlike traditional alloys, HEAs have promising properties such as crack resistance, excellent strength, ductility at low temperature, superior mechanical performance at high temperature, superconductivity, and high yield strength, making them an excellent candidate for structural material applications [2–9]. HEAs can be prepared by the arc melting process, powder metallurgy, and deposition technique among others. The alloy preparation process needs careful consideration of designing cost, complexity, and synthesis time to avoid the chances of getting undesirable mechanical properties. For example, HEAs such as AlNbTaTiVZr [10], AlHfNbTaTiZr [11], AlCrNbTiV [12], AlCrMoNbTi [13], CrNbTiZr [14], HfMoNbSiTiV [15], and HfNbTaTiZr [16] have high yield strength at room temperature but low yield strength at temperatures greater than 1200 °C, which significantly limits their application. Thus, an alternative computer simulation method is needed to predict the yield strength of HEAs at high temperatures before doing a trial-and-error experiment.

Recently, the machine learning (ML) method has been utilized to design the HEAs [17,18]. Islam et al. [19] developed a neural network (NN) model that can analyze the phase of multi-principal element alloys. Huang et al. [20] also used three ML techniques, support vector machine, artificial neural network, and K-nearest neighbors, to predict the phase of HEAs. Furthermore, ML techniques are used to predict the crystal structure of HEAs [21–23]. George et al. [24] also utilize the ML method for predicting the elastic properties of $Al_{0.3}CoCrFeNi$. Chang et al. [25] apply an artificial neural network to predict the hardness of AlCoCrFeMnNi. So far ML methods have proved the capability of predicting the properties of HEAs, which prompts us to predict the yield strength of HEAs at high temperature using ML. In the present study, we calculate the yield strengths of MoNbTaTiW and HfMoNbTaTiZr each at 800 °C, 1200 °C, and 1500 °C using the RF regressor model, and compare the predicted results with the available experimental data.

2. Computational methods

2.1. Data selection and analysis

Experimental data of yield strength of HEAs with compositions, testing temperature, and density were obtained from the report of Couzinie et al. [26]. To generalize the ML model, other features of HEAs such as mixing of entropy (ΔS_{mix}) [27], bulk modulus (B), valence electron concentration (VEC) [28], and melting temperature (T_m) are added in the datasets by using the following equations:

$$\Delta S_{mix} = -R \sum_{i=1}^{n} C_i ln C_i \quad (1)$$

$$B = \sum_{i}^{n} (C_i B_i) \quad (2)$$

* Corresponding authors.
E-mail addresses: uttam.bhandari@sus.edu (U. Bhandari), shizhong_yang@subr.edu (S. Yang).

https://doi.org/10.1016/j.mtcomm.2020.101871
Received 22 September 2020; Received in revised form 9 November 2020; Accepted 9 November 2020
Available online 13 November 2020

$$T_m = C_i\,(T_m)_i \tag{3}$$

$$VEC = \sum_{i=1}^{n} C_i (VEC)_i \tag{4}$$

where R represents universal gas constant, n represents the number of constituent elements, C_i, B_i, $(T_m)_i$, and $(VEC)_i$ represents the atomic percentages of the ith element, bulk modulus of the ith element, melting temperature of the ith element, and valence electron concentration of ith element, respectively.

Selection of irrelevant feature i.e. material properties in our case-during the ML training process can lower the model accuracy [29]. Therefore, it is essential to find out which feature is more crucial relative to the other features. To find the correlation between the features, we performed statistical analysis by calculating the Pearson correlation coefficient [30] using the Pandas library [31]. This correlation coefficient provides qualitative and quantitative interconnection between any two features of the dataset and is calculated using the following mathematical formula,

$$r_{xy} = \frac{1}{n-1} \frac{\sum_{i=1}^{n} (x_i - \bar{x}) - (y_i - \bar{y})}{\sigma_x \sigma_y} \tag{5}$$

where n is sample size, $\bar{x}$ and $\bar{y}$ are the mean values of two input features, σ_x and σ_y are the standard deviation of two features. The total value of correlation lies between -1 and 1, where 0 means no correlation, 1 means strong positive correlation, -1 means strong negative correlation. The correlation values are shown in the heat map diagram Fig. 1. It is seen from Fig.1 that there is a positive correlation between bulk modulus, melting temperature, and *VEC*. Testing temperature and yield strength have a negative correlation. However, the testing temperature has a positive correlation with entropy and density. In general, there is no strong positive and negative correlation matrix between any two features which indicates that all input features of datasets should be considered in the ML model.

Furthermore, we also plotted the scatter plots matrix as shown in Fig. 2, to understand the relationship between the two features. The histogram of diagonal panels shows that the yield strength prediction of HEAs depends on more than one input feature. It is also clear from Fig. 2 that bulk modulus, melting temperature, and *VEC* are very closely related as they produce a similar pattern of pair plot. Thus, the melting temperature and VEC were removed to make the learning algorithm faster and to avoid problems like multicollinearity, overfitting, etc.

2.2. RF regressor

RF regressor [32] is a popular technique in supervised learning, consisting of more than one decision tree for prediction. Each random tree prediction is either averaged or majority vote is considered to predict the output results [33]. The major benefits of the random tree are simplicity and ease of interpretation and understanding. The cross-validation method [25] from the Sklearn library is used to find the optimal values of *(i)* number of trees and *(ii)* maximum depth of trees in the RF regressor model. The general flowchart of the RF regressor model is shown in Fig. 3.

We ran the model on the dataset where the total number of features and samples are 25 and 238, respectively. Among them, 90% of the data was used to train the model and the remaining 10% was taken as the test data. We performed the cross validation 10 times to ensure that every data can participate in training and testing the model. In the model 100 trees or bootstrap samples with constant random features are prepared. Final predictions were done by taking the majority of votes among the predictions of all features. The performance of the model was evaluated by mean square error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \widehat{Y}_i\right)^2 \tag{6}$$

where n is number of data points, Y_i is observed values, $\widehat{Y}_i$ is predicted values.

3. Results and discussion

The predicted and experimental yield strength for present HEAs at 800 °C by RF model are presented in Fig. 4(a) and Fig. 4(b), which compares all training and testing data for the yield strength. All features including yield strength of training data were directed obtained from the dataset, so that all the dark blue points lie along the diagonal line. The predicted testing points which lie along the diagonal line are consistent with the measured value, which indicates that RF model is adequate for predicting the yield strength of HEAs. Indeed, the accuracy of RF model

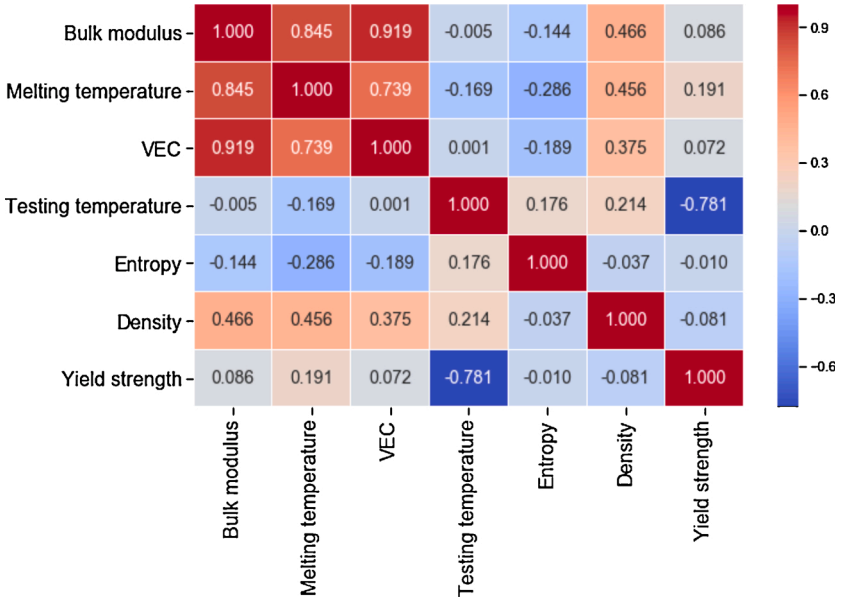

Fig. 1. Heat-map of the correlation matrix between the features present in the datasets.

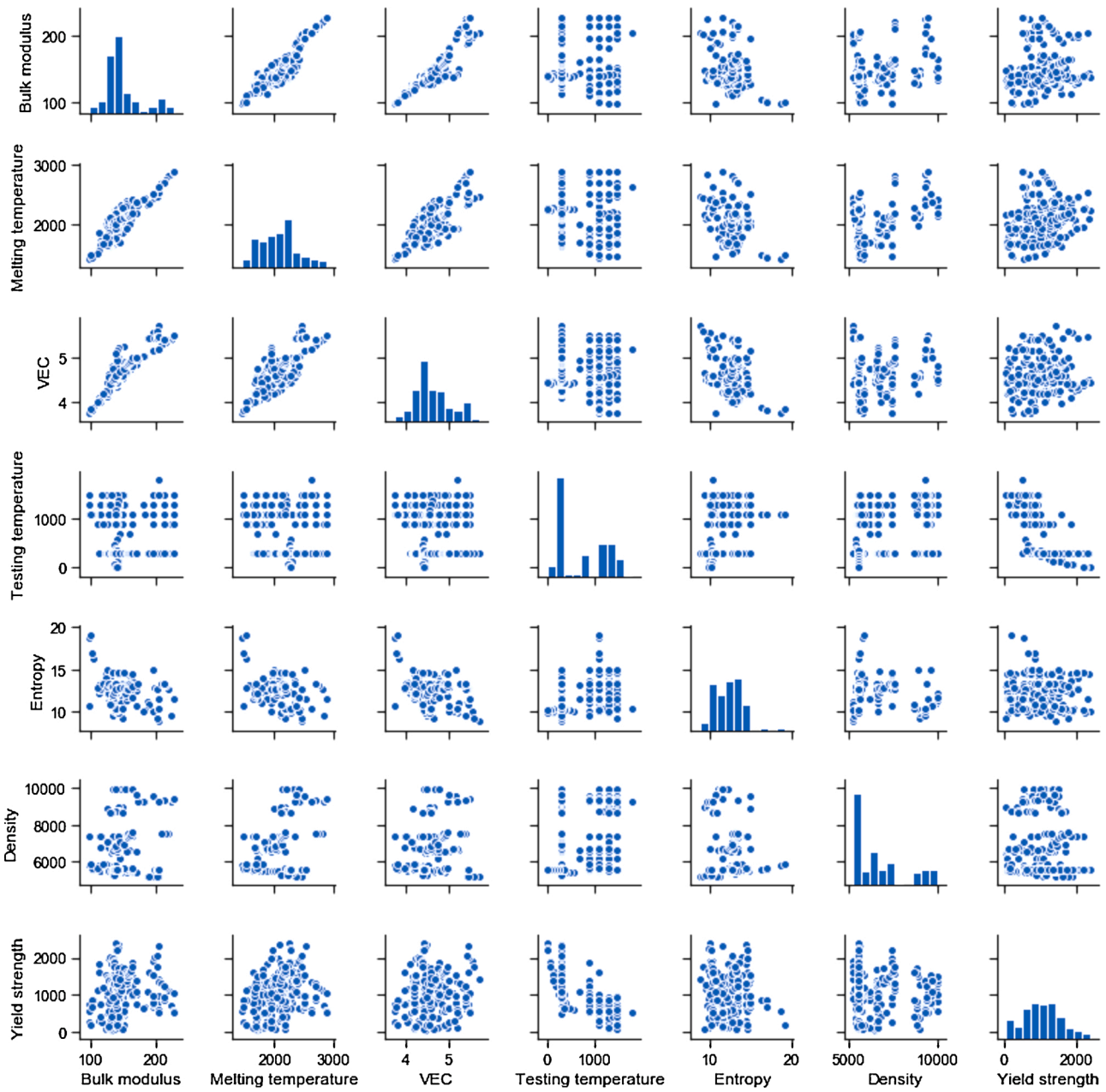

Fig. 2. Scatterplot matrices showing the relationships between all the features present in the 240 datasets.

in this work achieved 95% accuracy during training, which numerically demonstrates the high accuracy of the RF regressor model.

Based on the training results, the yield strength of MoNbTaTiW and HfMoNbTaTiZr each at 800 °C and 1200 °C, respectively were predicted. The elemental composition of present alloys and the value of the input features are supplied as inputs to get yield strength as an output. Table 1 shows the predicted value of yield strength by the RF regressor model. The accuracy percentage of MoNbTaTiW is higher than the HfMoNbTaTiZr. The lower accuracy of HfMoNbTaTiZr could be due to the lower number of datasets related to the alloy. In the future, improvement and addition of datasets will potentially increase the accuracy. The accurate prediction of yield strength by the RF regressor model confirms its ability to produce a good result even with smaller datasets.

After predicting the yield strength at 800 °C and 1200 °C accurately, we applied this model for predicting the yield strength at 1500 °C. Considering the accurate prediction of yield strength for current HEAs at 800 °C and 1200 °C, we expect the yield strength prediction at 1500 °C for current alloys will be confirmed by experiments in the future.

To check the importance of the input features, we removed one feature, i.e. density, while training the datasets and predict the yield strength of MoNbTaTiW and HfMoNbTaTiZr at 1200 °C. We found that the removal of the feature affects the testing accuracy. The predicted value of yield strength of MoNbTaTiW and HfMoNbTaTiZr is 492 MPa and 346 MPa, respectively. These yield strength values are lower than the experimental values. This signifies that all the selected input features in the sample data are important in predicting the yield strength of HEAs.

The classical ways of predicting the yield strength of high entropy alloys other than the experimental are first principles [36], rule of mixtures model [37], solid solution strengthening model [38], strengthening of screw dislocations theory [39], etc. Even though these approaches have been successful in predicting the yield strength, they possess their own limitations. Such computational methods need a very powerful CPU with more computational time which makes it more

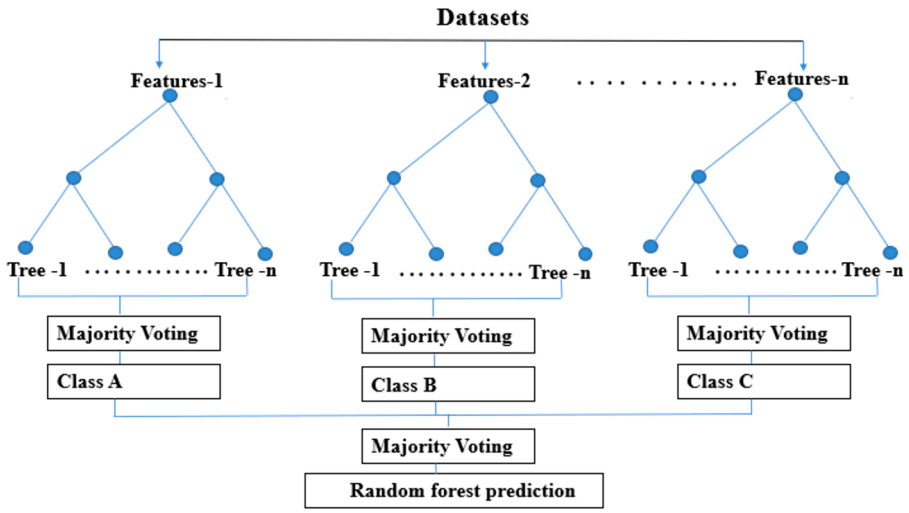

Fig. 3. Schematic diagram of random forest regressor model showing multiple decision trees with majority voting.

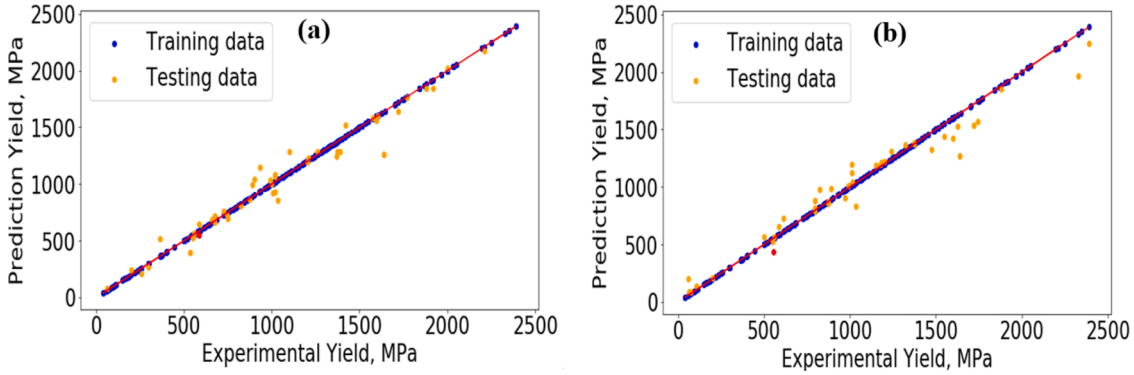

Fig. 4. RF regressor model prediction of yield strength using RF regressor model for (a) MoNbTaTiW (b) HfMoNbTaTiZr. The red point represents the yield strength prediction of HEAs at 800 °C. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 1
The yield strength (MPa) prediction of MoNbTaTiW and HfMoNbTaTiZr at 800 °C, 1200 °C, and 1500 °C by using the RF regressor model.

Alloys	Testing temperature (°C)	RF regressor prediction (MPa)	Experimental yield strength (MPa)	Error (%)
MoNbTaTiW	800	709	674 [34]	5.1
	1200	572	585 [34]	2.5
	1500	516	–	–
HfMoNbTaTiZr	800	929	1007 [35]	7.7
	1200	520	556 [35]	6.4
	1500	489	–	–

expensive to perform calculations. Suppose if we are required to design an HEA with desired yield strength, each alloy will have different yield strength for different composition of elements, then a trial-and-error method requires a lot of CPU time and designing costs. The basic steps of our ML algorithm are data collection, data preparation, training the model with the database, testing, and validating the model for prediction which is relatively simple and easy. We can change the composition of elements of HEA, calculate the required physical parameters, and predict the yield strength. The prediction of our model has been validated by comparing with the experimental report with an accuracy ranging between 93% and 97%. Furthermore, the ML method is mathematical based one which has advantages over classical methods on solving the problems where there are unknown physics and chemistry. The prediction accuracy of ML models was confirmed by comparing its results with the experimental report [19–25]. Our RF regressor model predicts the yield strength for two HEAs and the results are consistent with those obtained from the experiments, which shows that the model used has performed well. However, this model may fail to predict other mechanical properties such as hardness, Young's modulus, bulk modulus, and shear modulus as these properties may require different features of datasets. Moreover, the yield strength predicted by the model for HEAs may or may not be accurate, if the HEAs contain the element which is not present in the training data. Additionally, we believe that if the model is trained with sufficient data including extra features like grain size and precipitate effects, our model can predict the yield strength of future HEAs accurately. This approach will provide a guidance to design the virtual samples of HEAs until the desired yield strengths are reached. Similar ML method can be applied in predicting other mechanical properties such as hardness, stress, strength, etc. without any experiments in the future.

4. Conclusion

We have applied a ML method to predict the yield strength of HEAs at the desired temperature. This method is based on the RF regressor model that can predict the yield strength in less time and with low computational cost. The yield strength of MoNbTaTiW and HfMoNbTaTiZr each at 800 °C and 1200 °C, respectively, are predicted by the RF regressor model. The results obtained are in good agreement with the experimental data. This method will be very convenient in predicting the yield strength of HEAs at a temperature greater than 1200 °C, which,

otherwise, would be very difficult and expensive through the corresponding experiment.

Data and code availability

The data and code used in this study can be made available by the corresponding authors upon request.

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.

Acknowledgements

This research is partially supported by the NSFEPSCoRCIMM project under award No. 1541079, CIMM SURE awards No. LEQSF-EPS(2020)-SURE-237 and LEQSF-EPS(2020)-SURE-242, DOE award No. DE-NA0003979, and DoD support under contract No. W911NF1910005.

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.
[2] B. Gludovatz, A. Hohenwarter, D. Catoor, E.H. Chang, E.P. George, R.O. Ritchie, A fracture-resistant high-entropy alloy for cryogenic applications, Science 345 (6201) (2014) 1153–1158.
[3] U. Bhandari, C. Zhang, C. Zeng, S. Guo, S. Yang, Computational and experimental investigation of refractory high entropy alloy $Mo_{15}Nb_{20}Re_{15}15Ta_{30}W_{20}$, J. Mater. Res. Technol. 9 (4) (2020) 8929–8936.
[4] J. Miao, C.E. Slone, T.M. Smith, C. Niu, H. Bei, M. Ghazisaeidi, G.M. Pharr, M. J. Mills, The evolution of the deformation substructure in a Ni-Co-Cr equiatomic solid solution alloy, Acta Mater. 132 (2017) 35–48.
[5] Z. Li, D. Raabe, Strong and ductile non-equiatomic high-entropy alloys: design, processing, microstructure, and mechanical properties, Jom 69 (11) (2017) 2099–2106.
[6] Z. Li, S. Zhao, R.O. Ritchie, M.A. Meyers, Mechanical properties of high- entropy alloys with emphasis on face-centered cubic alloys, Prog. Mater. Sci. 102 (2019) 296–345.
[7] U. Bhandari, C. Zhang, S. Yang, Mechanical and thermal properties of low- density Al20+ xCr20-xMo20-yTi20V20+ y, Alloys Crystals 10 (4) (2020) 278.
[8] V.K. Soni, S. Sanyal, S.K. Sinha, Phase evolution and mechanical properties of novel FeCoNiCuMox high entropy alloys, Vacuum 174 (2020) 109173.
[9] U. Bhandari, C. Zhang, S. Guo, S. Yang, First-principles study on the mechanical and thermodynamic properties of MoNbTaTiW, Int. J. Miner. Metall. Mater. (2020) 1–7.
[10] O.N. Senkov, C. Woodward, D.B. Miracle, Microstructure and properties of aluminum-containing refractory high-entropy alloys, Jom 66 (10) (2014) 2030–2042.
[11] 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.
[12] N.D. Stepanov, N.Y. Yurchenko, D.V. Skibin, M.A. Tikhonovsky, G.A. Salishchev, Structure and mechanical properties of the AlCrxNbTiV (x= 0, 0.5, 1, 1.5) high entropy alloys, J. Alloys Compd. 652 (2015) 266–280.
[13] H. Chen, A. Kauffmann, B. Gorr, D. Schliephake, C. Seemüller, J.N. Wagner, H. J. Christ, M. Heilmaier, Microstructure and mechanical properties at elevated temperatures of a new Al-containing refractory high-entropy alloy Nb-Mo-Cr-Ti-Al, J. Alloys Compd. 661 (2016) 206–215.
[14] O.N. Senkov, S.V. Senkova, D.B. Miracle, C. Woodward, Mechanical properties of low-density, refractory multi-principal element alloys of the Cr–Nb–Ti–V–Zr system, Mater. Sci. Eng. A 565 (2013) 51–62.
[15] Y. Liu, Y. Zhang, H. Zhang, N. Wang, X. Chen, H. Zhang, Y. Li, Microstructure and mechanical properties of refractory $HfMo_{.50}NbTiV_{0.5}Si_x$ high-entropy composites, J. Alloys Compd. 694 (2017) 869–876.
[16] O.N. Senkov, J.M. Scott, S.V. Senkova, F. Meisenkothen, D.B. Miracle, C. F. Woodward, Microstructure and elevated temperature properties of a refractory TaNbHfZrTi alloy, J. Mater. Sci. 47 (9) (2012) 4062–4074.
[17] 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) 1–9.
[18] 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.
[19] N. Islam, W. Huang, H.L. Zhuang, Machine learning for phase selection in multi-principal element alloys, Comput. Mater. Sci. 150 (2018) 230–235.
[20] W. Huang, P. Martin, H.L. Zhuang, Machine-learning phase prediction of high-entropy alloys, Acta Mater. 169 (2019) 225–236.
[21] A. Agarwal, A.P. Rao, Artificial intelligence predicts body-centered-Cubic and face-centered-Cubic phases in high-entropy alloys, JOM 71 (10) (2019) 3424–3432.
[22] 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.
[23] 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) 1–9.
[24] G. Kim, H. Diao, C. Lee, A.T. Samaei, T. Phan, M. de Jong, K. An, D. Ma, P.K. Liaw, W. Chen, First-principles and machine learning predictions of elasticity in severely lattice-distorted high-entropy alloys with experimental validation, Acta Mater. 181 (2019) 124–138.
[25] Y.J. Chang, C.Y. Jui, W.J. Lee, A.C. Yeh, Prediction of the composition and hardness of high-entropy alloys by machine learning, JOM 71 (10) (2019) 3433–3442.
[26] J.P. Couzinie, O.N. Senkov, D.B. Miracle, G. Dirras, Comprehensive data compilation on the mechanical properties of refractory high-entropy alloys, Data Brief 21 (2018) 1622–1641.
[27] Y. Zhang, Y.J. Zhou, J.P. Lin, G.L. Chen, P.K. Liaw, Solid-solution phase formation rules for multi-component alloys, Adv. Eng. Mater. 10 (6) (2008) 534–538.
[28] S. Guo, C. Ng, J. Lu, C.T. Liu, Effect of valence electron concentration on stability of fcc or bcc phase in high entropy alloys, J. Appl. Phys. 109 (10) (2011) 103505.
[29] Y. Liu, T. Zhao, W. Ju, S. Shi, Materials discovery and design using machine learning, J. Mater. 3 (3) (2017) 159–177.
[30] J.D. Kelleher, B. Mac Namee, A. D'arcy, Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies, MIT press, 2015.
[31] W. McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, "O'Reilly Media, Inc.", 2012.
[32] Y. Lin, Y. Jeon, Random forests and adaptive nearest neighbors, J. Am. Stat. Assoc. 101 (474) (2006) 578–590.
[33] G.J. McLachlan, K.A. Do, C. Ambroise, Analyzing Microarray Gene Expression Data, 422, John Wiley & Sons, 2005.
[34] Z.D. Han, N. Chen, S.F. Zhao, L.W. Fan, G.N. Yang, Y. Shao, K.F. Yao, Effect of Ti additions on mechanical properties of NbMoTaW and VNbMoTaW refractory high entropy alloys, Intermetallics 84 (2017) 153–157.
[35] C.C. Juan, M.H. Tsai, C.W. Tsai, C.M. Lin, W.R. Wang, C.C. Yang, S.K. Chen, S. J. Lin, J.W. Yeh, Enhanced mechanical properties of HfMoTaTiZr and HfMoNbTaTiZr refractory high-entropy alloys, Intermetallics 62 (2015) 76–83.
[36] B. Yin, W.A. Curtin, First-principles-based prediction of yield strength in the RhIrPdPtNiCu high-entropy alloy, Npj Comput. Mater. 5 (1) (2019) 1–7.
[37] C. Varvenne, W.A. Curtin, Predicting yield strengths of noble metal high entropy alloys, Scr. Mater. 142 (2018) 92–95.
[38] H.W. Yao, J.W. Qiao, J.A. Hawk, H.F. Zhou, M.W. Chen, M.C. Gao, Mechanical properties of refractory high-entropy alloys: experiments and modeling, J. Alloys Compd. 696 (2017) 1139–1150.
[39] F. Maresca, W.A. Curtin, Theory of screw dislocation strengthening in random BCC alloys from dilute to "High-Entropy" alloys, Acta Mater. 182 (2020) 144–162.