



This is a repository copy of *Supervised machine learning for multi-principal element alloy structural design*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/218181/>

Version: Published Version

Article:

Berry, J. orcid.org/0000-0001-7291-2306 and Christofidou, K.A. orcid.org/0000-0002-8064-5874 (2024) Supervised machine learning for multi-principal element alloy structural design. *Materials Science and Technology*. ISSN 0267-0836

<https://doi.org/10.1177/02670836241272086>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Supervised machine learning for multi-principal element alloy structural design

Joshua Berry¹  and Katerina A. Christofidou¹ 

Materials Science and Technology

1–18

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/02670836241272086

journals.sagepub.com/home/mst



Abstract

The application of supervised Machine Learning (ML) in material science, especially towards the design of structural Multi-Principal Element Alloys (MPEAs) has rapidly accelerated over the past five years. However, several factors are limiting the impact that these ML methodologies can have, chief amongst them being the availability and fidelity of data. This review analyses how ML has been utilised to accelerate the design of novel structural MPEAs, outlining the standard procedures followed, and highlighting the successes and common pitfalls identified in current studies. The need for experimental validation and incorporation into closed loop ML pipelines is also discussed, including the influence and integration of manufacturing methodologies into the ML decision making process.

Keywords

multi-principal element alloys, high entropy alloys, complex concentrated alloys, compositionally complex alloys, machine learning, experimental data, alloy design

Received: 2 February 2024; accepted: 2 July 2024

Introduction

High Entropy Alloys (HEAs), first introduced to the scientific community in 2004 by Yeh et al.¹ and Cantor et al.² respectively, are conventionally defined as a class of alloys containing five or more elements in either equiatomic elemental concentrations, or elemental concentrations in the range of 5 to 35 at.%.¹ This concept leads to HEAs occupying a vast uncharted compositional space³ and sparking a wealth of studies and debates in the literature, not least on appropriate naming conventions. Consequently, several different terms have been proposed and are used to encompass different classes of materials such as, multi-component alloys, compositionally complex alloys, complex concentrated alloys or indeed the broader term, Multi-Principal Element Alloys (MPEAs). Concurrently, the term HEA has evolved to more routinely describe single phase MPEAs.^{4,5} For consistency in this review, MPEA will be used to refer to all subclasses listed above. Recent review articles and critical assessments of the MPEA field are available for readers unfamiliar with the background and application of these materials.^{4–8}

Machine Learning (ML) is a methodology whereby computer systems can learn to perform specific data-based tasks without any explicit programming. Broadly speaking, ML can be split into three different categories, supervised, unsupervised and reinforcement learning.⁹ This review will focus on the application of supervised learning, where ML models are fit on data containing known target outputs.

Hence, the model can be trained to recognise patterns and trends in the available input data to predict the output.¹⁰ Supervised learning can be further subdivided into classification and regression tasks. Classification is used to categorise the discrete values of a variable and separate the predictions into different categories. In contrast, regression is used to predict continuous numerical values.⁹

With microstructural simplicity being a founding principle of the field, efforts have been made to develop rules to enable the prediction of the formation of solid-solution phases. These rules commonly take the form of two-dimensional phase stability plots and are based on the Hume-Rothery, Gibbs free energy and valence electron criteria.^{11–18} Furthermore, a significant number of experimental studies have been conducted to determine the microstructural and mechanical properties of different MPEA compositions, fabricated through a variety of manufacturing techniques as well as evaluating the impact of post processing methods.^{19–24} The generation of rules describing the phase stability of MPEAs, continuous experimental data collection, and vast compositional space, results in a

¹Department of Materials Science and Engineering, The University of Sheffield, Sheffield, UK

Corresponding author:

Joshua Berry, Department of Materials Science and Engineering, The University of Sheffield, Sheffield, UK.

Email: jberryl@sheffield.ac.uk

natural relationship forming between ML and the MPEA field. As ML tools and materials data have become more accessible, especially since 2014, the application of ML in the field of material science has been growing exponentially.²⁵ ML is rapidly becoming an accepted tool to automate materials discovery,^{26,27} expediting searches of the compositional space in an unconstrained manner,²⁸ whilst minimising the need for expensive trial and error experiments.²⁹ ML can guide these experimental investigations by reviewing large amounts of data to discover patterns and trends in higher dimensions than possible for humans. Subsequently enabling downselection of compositions with desirable mechanical and microstructural properties for the intended applications.^{26,29} These predictions can be performed quickly and produce informative results while being reproducible with the capability for future scaling.⁹ Furthermore, newly available data can be directly incorporated into future iterations to improve the prediction capabilities.²⁸

However, despite these advantages there are several drawbacks to the application of ML to the MPEA field. Firstly, the success and development of ML within the MPEA field is intrinsically linked to the experimental exploration of the compositional space.²⁶ Diverse and expansive databases are required to train supervised ML algorithms, but MPEA databases typically only contain a few hundred to a few thousand data points.^{29,30} ML calculations are also susceptible to overfitting, where an ML model too closely matches the training data and may be unable to make more generalised predictions on unseen data.²⁷ Furthermore, ML models are typically seen as “black boxes”, with limited interpretability into the internal mechanisms that map the input features to the target outputs⁹ and a lack of materials science insights. Interpretability refers to the transparency of the model’s decision making process and how easy the methodology of decisions is understood.³¹

Therefore, it is the aim of this review to investigate the implementation of supervised ML for the design of MPEAs through physical property predictions and the challenges associated with utilising this methodology. The order of this review follows the key steps of a typical ML study, Figure 1. Firstly, the availability of MPEA data suitable for ML is discussed. Secondly, the different input features and ML algorithms utilised in recent studies on ML for structural MPEA design are assessed. Finally, experimental validation of ML predictions and the impact of manufacturing on the application of ML within the MPEA field is examined and areas of further work identified and discussed.

Multi-principal element alloy experimental data

In any supervised ML task, the first step is collecting the appropriate data with which to train and test the models,³¹ as shown in Figure 1. ML performs best when the data and its interrelations are too complex for humans to

rationalise and often fails to produce meaningful relationships from small amounts of data.³⁰ Hence, to maximise the ML models’ performance and predictive power, large volumes of high-quality data are required.^{26,32–34} Furthermore, the size of the available data is critical in determining the optimal ML algorithm for the task. For small datasets, classical and statistical models such as regression, clustering and tree based methods perform best.³⁰ Significantly more complex algorithms, such as neural networks, require large amounts of data, thus, are only suitable when thousands of unique datapoints are available.^{30,35}

Despite significant research and publications in the field of MPEAs since 2004, the available MPEA data for ML tasks is still sparse.^{33,36,37} Many ML studies within the MPEA field report that this paucity of data is the main limiting factor in the performance of ML models.^{27,32,36,38–45} In addition, MPEA data is highly imbalanced because of the initial emphasis towards the discovery of single-phase solid solution microstructures and tendency to not report negative results in peer-reviewed literature.^{3,5,10,25,35} When dealing with real world data, duplicated, poorly formatted, and irrelevant and incomplete data is an unavoidable problem. Therefore, to prepare the raw data for analysis, data cleaning and pre-processing is a critical, yet often undocumented step within MPEA ML studies.^{30,31,33,46}

Table 1 summarises a series of experimental MPEA databases available in the published literature, combinations of which are often used as the foundational training data for many published ML studies.⁴⁷ Table 1 highlights that significantly more work is needed in making well populated MPEA data open access in standardised formats, to prevent data siloing and facilitate data sharing.^{25,26,48} All reported results and data from MPEA focussed ML studies should be provided, but this is very often not the case.³⁰ In addition to these experimental databases, there are many databases constructed from modelling simulations. One example of this is the Materials Project, providing structural and property information on over 150,000 inorganic materials, including MPEAs, from Density Functional Theory (DFT) calculations.⁴⁹

Because of the relative lack of availability of MPEA data, ML models are often constructed using sparse datasets. For example, to the authors’ knowledge, the largest experimental training dataset was used by Pei et al.⁵⁹ and consisted of 1252 datapoints. Conversely, a significant number of studies constructed models from databases containing 200 or fewer datapoints.^{36,40,42,60,61}

To combat the insufficient data within the MPEA field, many data supplementation techniques have been utilised. The CALculation of PHase Diagrams (CALPHAD) method has been utilised to both generate databases and supplement existing experimental data. Vazquez et al.³⁵ performed single-point equilibrium calculations to generate an MPEA database of 229,156 compositions. Similarly to CALPHAD, DFT calculations have also been applied to generate MPEA databases for ML training. Zhang et al.³⁷ generated a dataset of 3579 quaternary MPEAs while

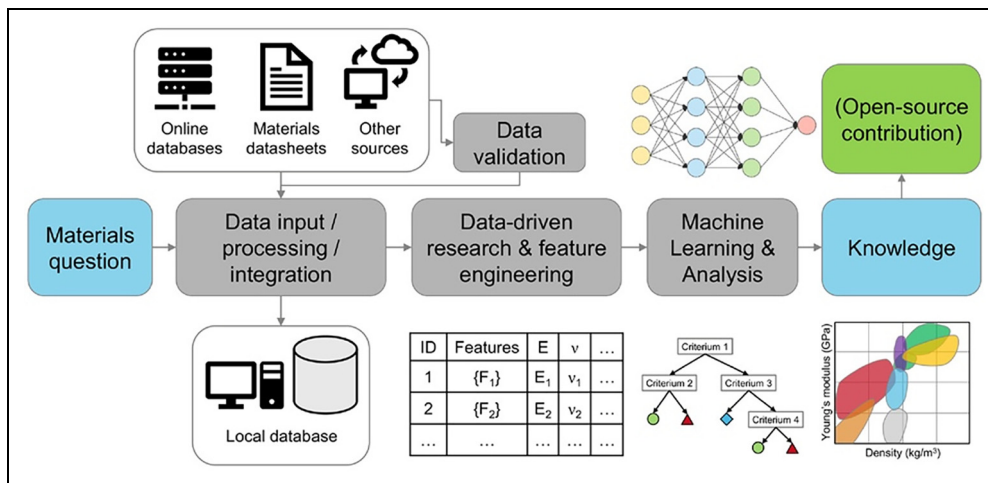


Figure 1. Example of a typical ML study workflow or pipeline within materials science. Reprinted with permission from.³⁰ ©2024 American Chemical Society

Kaufmann et al.⁴³ supplemented 134 experimental compositions with 1664 DFT compositions. This use of CALPHAD and DFT to generate data for ML training raises the question, if it is possible to generate data in this way to train ML models, why use ML? There are multiple reasons for this. Firstly, CALPHAD is calculated based on energy minimisation of experimental data, which for most alloys, often requires extrapolation from binary and ternary systems that comprise the thermodynamic databases. Secondly, ML is significantly faster and incurs lower computational costs³⁸ than both CALPHAD and DFT calculations. Vazquez et al.³⁵ reported that their neural network is 436 times faster than the CALPHAD method, and Kaufmann et al.²⁸ stated that DFT can take 100s of hours of computation per composition.

As illustrated in Table 1, MPEA data on properties such as phase, collected under a variety of experimental conditions, are published in small datasets across a range of literature. Therefore, it would seem logical to aggregate these data sources to enhance the quantity of MPEA data for ML model training. However, Ottomano et al.⁶² argue that expanding datasets in this way may affect the organicity and overall quality of the available data. To demonstrate this, model performance of a range of algorithms was compared before and after aggregation using a variety of collation techniques. In all cases, classical ML algorithms such as random forest and logistic regression demonstrated a reduction in performance. In contrast, deep learning models showed greater robustness, but no significant change in performance. Therefore, when collating data to increase data quantity, it is critical to consider the type of ML algorithm being utilised and the difference in origin of the individual data.

Other methods to combat insufficient data can be as simple as employing an empirical relationship. For example, Equation 1 was used by Huang et al.²⁷ to convert reported yield strength data σ_y , to the target variable hardness HV . However, such practices come at the risk of introducing low quality data into the training dataset.

Equation 1 for example, has been found to be accurate for BCC MPEAs, but less so for FCC MPEAs.²⁷ Thus, such practices are generally not recommended for MPEA studies, unless sufficient evidence is present to ensure the fidelity of generated data. Alternatively to physics based or empirical modelling approaches towards data generation, ML can also be used to generate data for ancillary ML algorithms. Lee et al.⁶³ produced a Conditional Generative Adversarial Network (CGAN) to generate additional training data, and demonstrated that employing a CGAN in conjunction with a neural network can improve phase prediction performance for MPEAs. However, the number and diversity of the generated samples depends on the original dataset. Hence, on smaller datasets, CGAN augmentation has limited impact on model performance.³² Finally, Paliania et al.¹⁰ discusses the potential shift in focus towards utilising natural language processing as a technique to strip materials knowledge and data from published literature. This could efficiently compile existing materials knowledge to produce substantial databases for MPEA based ML studies.^{64,65}

$$\sigma_y(MPa) \approx \frac{1}{3}H(MPa) = \frac{9.81}{3}HV(Hv) \quad (1)$$

Tackling the imbalance of MPEA data, highlighted in Figures 2 and 3, is an entirely different problem. Traditional classification tasks struggle with imbalanced data as they have a tendency to categorise data into the majority class,^{31,39} which also tends not to be the class of interest in the majority of studies. For small imbalanced datasets the minority class is typically insufficient for learning, especially when there is a degree of overlap within the classes.³¹ This is especially true for MPEA databases, where FCC and BCC phases dominate, and pronounced overlap exists between solid-solution and solid-solution plus intermetallic phases.⁶

The most common methods to tackle the issue of imbalanced data are random oversampling and undersampling.³¹ Oversampling supplements the available training data with

Table 1. Experimentally calculated MPEA databases published in peer-reviewed literature.

Authors	Year	Material Properties Included	Processing Method & History Included	Number of Datapoints
Ye et al. ⁵⁰	2016	Phase	All As-Cast	118
Miracle et al. ⁶	2017	Phase as SS or IM. Material properties are mentioned but not explicitly included.	Yes (Processing and post-processing)	648
Gorsse et al.* ^{51,52}	2018	Phase, ρ , HV, σ^y , σ^{UTS} , ϵ , E	N/A	370
Couzinié et al. ⁵³	2018	Phase, ρ , σ^y , E	Yes (Processing)	340
Roy et al. ⁵⁴	2020	Phase, (E)	N/A	340, (107)
Borg et al. ⁵⁵	2020	Phase, ρ , HV, σ^y , σ^{UTS} , ϵ , E	Yes (Processing)	1545
Machaka et al. ⁵⁶	2021	Phase	Yes (Processing and post-processing)	1360
Detor et al. ⁵⁷	2022	HV	Yes (Processing)	86
Han et al. ³⁴	2022	Phase	Yes (Processing, all as-cast)	1138
Chen et al. ⁵⁸	2022	Phase, Strength, Low-Cycle Fatigue, High-Cycle Fatigue and Fatigue Crack Growth Rate	Extensive processing methodology and history	66

ρ denotes the density, HV represents the Vickers hardness, σ^y is the yield strength, σ^{UTS} is the ultimate tensile strength, ϵ is the strain elongation and E represents the Young's modulus.

*Corrigendum

clones of the minority classes in order to balance them with the majority class.⁶⁶ Risal et al.⁶⁶ applied random oversampling and found it to be effective for phase prediction based on the available dataset. Ren et al.³³ also utilised random oversampling to change the distribution of their datasets and although the authors did not directly comment on the impact of the data augmentation, they successfully produced two models to predict high hardness MPEAs. In contrast, undersampling modifies the class distribution by reducing the data of every class to match the minority class.⁶⁶ This technique is significantly less popular in the MPEA field as it means the removal of precious data from already small datasets. Hence, oversampling is generally preferred. Another technique to combat data imbalance is Synthetic Minority Oversampling Technique (SMOTE).⁶⁷ SMOTE generates synthetic samples based upon the minority classes of the input dataset, using k-nearest neighbours to sample the feature space of the class, such that the distribution of data across all the classes is balanced. Crucially, this method increases the number of datapoints while maintaining the original trend.^{39,68} Bansal et al.⁶⁸ reported an improvement in model prediction by increasing the quantity of datapoints using synthetic data. However, Singh et al.⁶⁹ argue that implementing augmented data to combat data imbalance is not reliable. The authors claim that accuracy alone is not the most robust measure for assessing performance from ML models constructed from imbalanced data and, that it cannot be guaranteed the generated samples are MPEAs. To investigate this, multiple “vanilla” ML classifier models were compared with SMOTE-Tomek augmented models of the same algorithms. The best performing classifier was found to be a random forest model and despite the claims that augmenting data is not the optimal approach the augmented models showed better performance scores in all cases than the vanilla models. Hareharen et al.⁷⁰ also find that SMOTE improves the

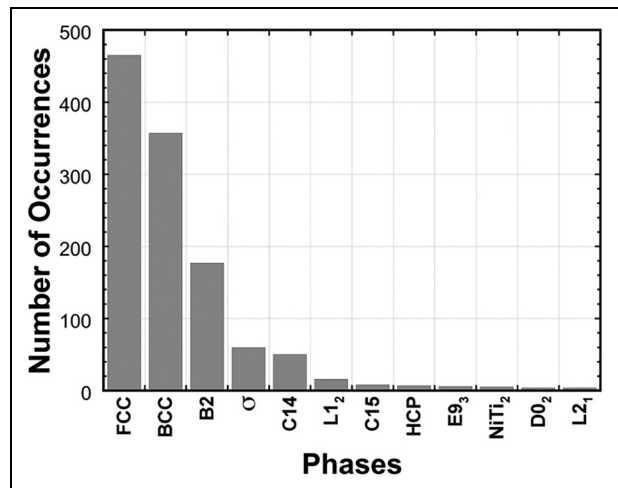


Figure 2. Data imbalance and domination of FCC and BCC phases across 648 published reports of microstructure within MPEA studies. Phases appearing less than 4 times are not shown.⁶ ©2024 reused with permission from Elsevier

ML model's ability to differentiate between the various classes.

Multi-principal element alloy feature selection

Supervised ML tasks aim to construct models to predict a target variable from a set of input features.⁹ In the case of MPEAs, the target is often phase formation^{47,71} as a proxy for microstructure, which dictates structural and mechanical properties. Thus, the input features are most commonly empirical relations based on the atomic properties of the alloys' constituent elements describing electronic, thermodynamic, physical and chemical characteristics.^{32,72} Table 2 summarises several examples of features frequently utilised in MPEA ML models.

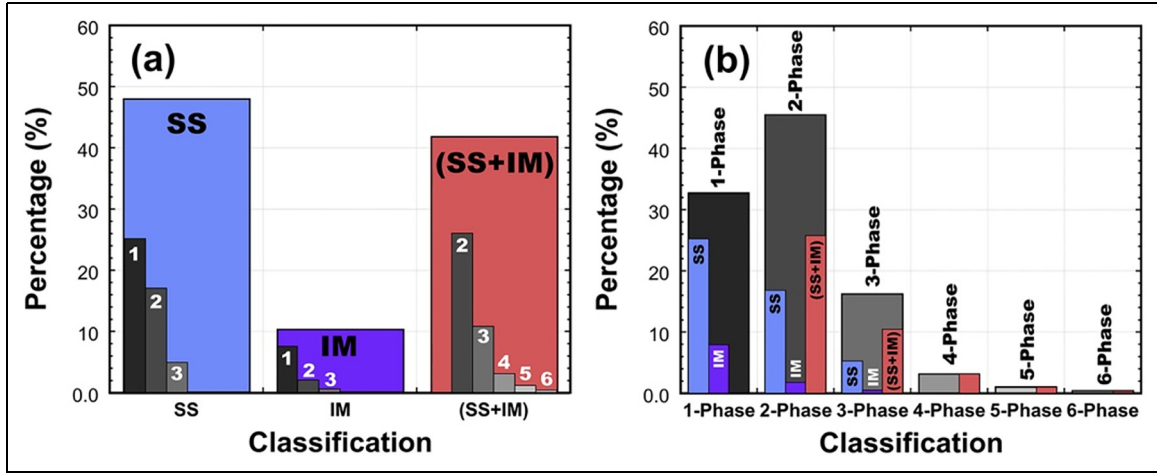


Figure 3. Classification of solid solution and intermetallic phases within MPEA data. (a) Microstructure classification by phase type with sub-classification by number of phases, (b) Number of phases classification with sub-classification by type of phase.⁶ ©2024 reused with permission from Elsevier

Table 2. Examples of features commonly used in ML models in the MPEA field

Feature	Name	Equation	Reference
δ	Atomic Size Difference	$\delta = 100 \sqrt{\sum_{i=1}^n c_i \left(1 - \frac{r_i}{\bar{r}}\right)^2}$	18,73,74
γ	Atomic Packing Parameter	$\gamma = \frac{\omega_S}{\omega_L}, \quad \omega_x = 1 - \sqrt{\frac{(r_x + \bar{r})^2 - \bar{r}^2}{(r_x + \bar{r})^2}}$	18
$\Delta\chi$	Electronegativity Difference	$\Delta\chi = \sqrt{\sum_{i=1}^n c_i (\chi_i - \bar{\chi})^2}$	75
VEC	Valence Electron Concentration	$VEC = \sum_{i=1}^n c_i (VEC)_i$	13,76
e/a	Number of Itinerant Electrons per Atom	$\frac{e}{a} = \sum_{i=1}^n c_i \left(\frac{e}{a}\right)_i$	13,76,77
ΔH_{mix}	Enthalpy of Mixing	$\Delta H_{mix} = 4 \sum_{i=1, i < j}^n c_i c_j \Omega_{ij}$	11,14,78
ΔS_{mix}	Entropy of Mixing	$\Delta S_{mix} = -R \sum_{i=1}^n c_i \ln c_i$	8,14
T_m	Weighted Melting Temperature	$T_m = \sum_{i=1}^n c_i (T_m)_i$	12
Ω	Yang Parameter	$\Omega = \frac{T_m \Delta S_{mix}}{ \Delta H_{mix} }$	12
$\Delta\epsilon$	Young's Modulus Asymmetry	$\epsilon = \sqrt{\sum_{i=1}^n c_i \left(1 - \frac{E_i}{\bar{E}}\right)^2}$	79

Quantities denoted with a bar indicate that it is the average value, while quantities with an i indicate the i 'th element. c represents the atomic fraction of the element, r denotes the atomic radii and E represents the Young's modulus. R is the molar gas constant. For γ , ω_x denotes the solid angles around the largest and smallest atoms, represented by subscript L and S respectively. In ΔH_{mix} , Ω_{ij} is the enthalpy coefficient for elements i and j respectively.

Features, such as those summarised in Table 2, often originate from the Hume-Rothery rules on solid-solution formation for binary systems. For example, the need for small atomic size differences, comparable valency and similar electronegativities.⁸⁰ These features have been shown to display clear trends with MPEA microstructural and mechanical properties, thus, successfully translating to MPEA ML studies. Guo et al.¹³ determined Valence Electron Concentration (VEC) to be the physical parameter controlling formation of FCC ($VEC \geq 8$) or BCC ($VEC < 6.87$) solid solutions. This observed phase formation has a

strong impact on mechanical properties with BCC phases typically observed to display a higher hardness,⁸¹ but a lower ductility compared to their FCC counterparts.^{8,82,83} Similarly, Wang et al.¹⁸ developed a new parameter, γ to describe atomic packing as an improvement to the commonly accepted atomic size difference, δ , where $\gamma < 1.175$ results in solid-solution formation. Indeed, these trends have also been observed when applied to ML studies, with VEC and δ being reported as two of the most important features by many authors investigating both microstructural and mechanical properties,^{27,33,41,42,84} demonstrated in

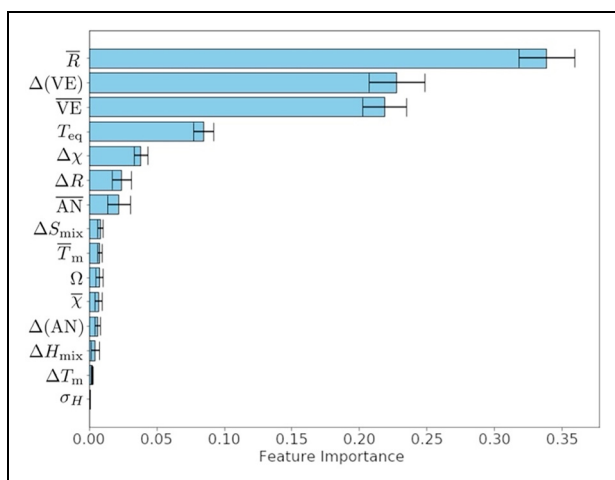


Figure 4. Feature importance scores of 15 features within an XGBoost model used for phase classification. In this case R denotes atomic radius and VE represents valence electron count.⁸⁵ ©2024 adapted with permission from Elsevier

Figure 4. This importance is defined and measured according to how much impact the feature has on the models decision making process.

An important decision when selecting features for MPEA ML studies is whether to include the chemical composition of the alloy as an input feature. Bakr et al.⁴⁵ demonstrated that it is possible to train an artificial neural network for phase and hardness prediction of MPEAs using only the chemical composition as the input feature. Similarly, Jain et al.⁴⁶ included the elemental compositions as features for phase predictions and used them as the only feature for hardness predictions. However, Wen et al.⁸⁶ argue that utilising the elemental composition, in conjunction with the elemental property features discussed above, significantly out-performs just training ML models on the elemental composition. Furthermore, Morgan et al.²⁵ state that utilising composition as an input feature means that the model cannot be used to extrapolate to systems including elements outside of the training data. Hence, they argue that it is better to represent each element with elemental properties to enable feature generation by taking compositionally averaged combinations of the constituent elements, as in Table 2. If utilising ML to optimise the composition of an already defined MPEA system, then including composition may be useful. For example, Chen et al.⁴¹ include the molar fraction of the six constituent elements of their MPEA system as features for the model, but only hardness optimisation of a single alloy system was investigated. Alternatively, if the goal of ML application is materials discovery, then evidence suggests prioritising elemental property features.

Subsequent to data collection, feature engineering is the next step in model construction and consists of feature generation and selection. Firstly, the features chosen for any ML model must be both machine readable and relevant to the target variable.²⁵ Following the discussion of the need for high quantities of data to train ML models, it could be

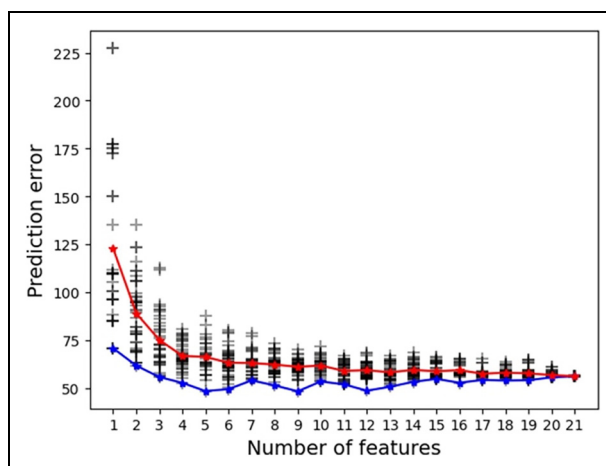


Figure 5. Model performance plateaus as the number of features is increased, in this case for a support vector machine model. The red and blue lines represent the average and lowest prediction error for that number of features respectively.⁸⁸ ©2024 reused with permission from Elsevier

assumed that the more features available to make predictions of the target, the better a model will perform. However, this is not the case and many MPEA ML studies demonstrate that as the number of features increases, prediction accuracy plateaus,^{33,86–88} as illustrated in Figure 5. This also introduces feature redundancy, overfitting and results in poor generalisability and computational efficiency.^{25,33,88} Feature selection is an essential process that reduces the dimensionality of the ML task by identifying and removing irrelevant, noisy and redundant features. Simultaneously feature selection retains sufficient information and enables optimisation of the number and combination of features to maximise accuracy and improve ease of training.^{31–34,38,47,89}

Collinearity of features is detrimental to model performance,³³ computational efficiency and, crucially, interpretability. Further, it restricts the ability to ascertain the individual contribution of each feature to the model.⁹⁰ Hence, for features that are correlated, the least important feature is typically omitted.^{40,44} To detect collinearity, the most popular method is to measure the correlation of individual features using the Pearson Correlation Coefficient (PCC),^{34,91} given by Equation 2.

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

Where x and y denote two of the features and \bar{x} and \bar{y} represent the mean of the two features, respectively. PCC values can range from +1 to -1, with positive values indicating a positive relationship between the variables and vice versa. Commonly in correlation analysis, features with PCC values > 0.80 are considered very strongly correlated.^{33,41,91} However, there are no fixed cut-offs for the interpretation of feature correlation strength, rather context is critical to understanding both the extent and impact of the correlation on the model.^{92,93} Despite this, in many MPEA ML studies, the limit for where features are considered highly correlated

is often set very close to 1, at $PCC > 0.95$.^{40,84,88} Typically, in MPEA ML studies where the correlations are moderate, the issue of where to place the limit of correlation is avoided and authors often do not state at what point they would have considered features to be correlated and why.^{42,44,72} It is also common for MPEA ML studies to assess how important ML features are through their correlation with the outputs,⁴⁶ but this is not an effective method for model interpretability.

Data, features, and models are extensively interlinked and so the optimal feature combinations for the available data must be found for each model tested. However, it can be computationally prohibitive to test every possible permutation.²⁵ Several techniques and approaches have been developed to enable feature engineering. The first and simplest method of feature selection is human context. Nonsensical features and those with no relation to the target variable should be removed.²⁵ Sequential feature selection iteratively adds or subtracts features from the dataset in order to maximise model performance.^{25,33,87} Li et al.⁸⁸ utilised a genetic algorithm to find the optimal combination of features. The genetic algorithm mimics the mechanism of natural selection to arrive at the global optimum performance without testing every possible combination. This saves on computational efficiency over the exhaustive sequential feature selection methods, especially for large feature spaces. MPEA ML studies mostly agree that fewer features (4 or 5) are optimal for predicting phase formation^{33,86–88} and mechanical properties.²⁹ However, although Huang et al.⁴⁰ agrees that the optimal number of features for phase selection is small (5), they find that the optimal number of features for hardness prediction is much larger (13), in contrast to the majority of MPEA studies.

Machine learning algorithms for multi-principal element alloy design

An exhaustive discussion on how each individual supervised ML algorithm functions is beyond the scope of this review. See Hastie et al.⁷¹ and James et al.⁹⁰ for an overview. No supervised ML algorithm is by default superior to any other and the choice of optimal algorithm depends strongly on the available data and target output.^{30,87} Choosing a suitable algorithm is critical for improving model performance and efficiency.³³ Therefore, many ML studies trial a handful of potential ML algorithm candidates to determine which performs the best, for example,^{34,40,87} shown in Figure 6. Alternatively, some ML MPEA studies already have a specific algorithm in mind because of the benefits of that particular methodology. For example, Bhandari et al.⁴⁴ and Choudhury et al.⁹⁴ select a random forest algorithm for its simplicity and interpretability.

Simple linear regression models are much easier to interpret but lack in predictive power and performance.³⁰ Clustering algorithms, such as k-nearest neighbours, often suffer from low accuracy and when the data is imbalanced,

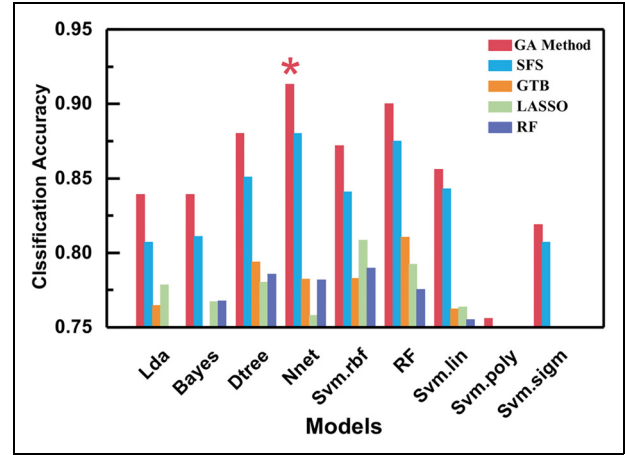


Figure 6. Comparison of model performance across the same experimental database. Each column represents a different feature subset selected by application of the models indicated in the legend.⁸⁷ ©2024 adapted with permission from Elsevier

the more frequent classes significantly dominate predictions, becoming more noticeable as the number of neighbours considered increases.^{27,34} Tree based algorithms have very high interpretability and training speed but are susceptible to overfitting. However, ensemble tree based models such as random forest have anti-overfitting properties.^{10,34} In contrast, support vector machines and neural networks can provide a boost in predictive performance at the cost of model interpretability and the need to scale input data. Lastly, neural networks require significantly more computational power and training time for small gains in performance.^{10,30,31} Figure 7 provides an illustration of this transparency to performance trade-off.

To assess the predictive performance of supervised ML models a range of metrics are available, split between the classification and regression tasks, that assess the difference between the model's predicted outputs and the true outputs from the validation data set.²⁵ For classification tasks, accuracy, precision, recall, and F1-score are the most common performance metrics in MPEA ML studies, provided in Equation 3.^{39,45,47,94} A complete description of these can be found in Dalianis.⁹⁵

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3a)$$

$$Precision = \frac{TP}{TP + FP} \quad (3b)$$

$$Recall = \frac{TP}{TP + FN} \quad (3c)$$

$$F_1 \text{ Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3d)$$

True positives (TP) are the data points that are correctly predicted by the model for a class. False negatives (FN) are data points that are incorrectly predicted as a different class by the model. False positives (FP) are data points

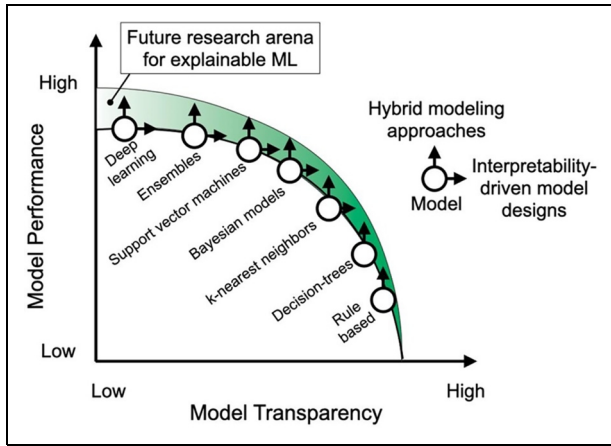


Figure 7. Trade-off between model transparency and model predictive performance.¹⁰ ©2024 reused with permission from Elsevier

that are a different class but are predicted to be the class under consideration. True negatives (TN) are data points that are a different class to the one under consideration and are predicted to be a different class, thus the class being considered is not involved. For regression models, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are the two most common metrics that MPEA ML models will aim to minimise, Equation 4a and 4b respectively.^{25,32,33,38,96} In addition, the coefficient of determination (R^2), Equation 4c, combined with parity plots of predicted vs actual data, provides an informative assessment of model quality, with values >0.7 typically indicating a useful model.²⁵

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4a)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (4b)$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (4c)$$

Where y denotes the true value, \hat{y} represents the predicted value, \bar{y} represents the mean of the true data, and n is the number of data points. To enable the model performance to be correctly validated using the metrics described above, the available data first needs to be split into training, validation and testing in a reproducible manner.³⁰ The training data is used by the model to relate the input features to the desired output. The validation data is then used to assess the performance of the model on the training data and optimise the chosen algorithm. Subsequently, the testing data is used to assess the error in the final optimised version of the model.^{25,30}

There are two common ways to split the data to train an ML model and assess its performance. The first is Train-Test-Split where the database is simply split into fixed fractions (e.g., 80% – 10% – 10%). However, this can cause poor generalisability in the data across the

fractions if the data distribution is imbalanced.³³ Bakr et al.⁴⁵ reported a significant difference between the RMSE of the training and testing data due to a different distribution of data across the two sets. The second approach is k -fold cross-validation, Figure 8. In this method the dataset is divided into k equal fractions, with each fraction used once as validation data while the other $k-1$ fractions are used for training. After k iterations, the model will have been trained and validated k times on the same database, eliminating the effect of sampling. This enables the average performance metrics to be calculated, providing a more appropriate assessment of model performance, and is particularly useful for small datasets, such as those in the MPEA field.^{27,33,45} However, while the majority of MPEA ML studies detail the methodology they have used to assess the performance of the model on the available data, very few discuss retraining the final model using the whole database. This is a key step to maximise the amount of training data. As it is not often discussed in the published literature, it is unclear if this step is performed in MPEA ML studies, with the notable exception of Liu et al.³⁶

Hyperparameter optimisation is a key step in the development of ML models and features in the vast majority of MPEA ML studies. Hyperparameters are guidelines for the ML model construction that control the learning process.³⁰ For example, they could be controlling the maximum number of trees and branches in a random forest or the number of nodes and layers in a neural network.⁷² It is important to optimise the hyperparameters of an ML model to maximise the model performance and speed.^{33,96} The two most common methods of hyperparameter optimisation within MPEA ML studies are grid search and Bayesian optimisation. For grid search, a range for each hyperparameter is specified and each possible combination of these parameters is tested against a baseline to find the optimal combination. This can also be performed randomly rather than exhaustively to save on computation time. In contrast, Bayesian optimisation minimises the loss on an error function to find the optimal hyperparameters for a model. Neither method is by definition better than the other, but both have been used successfully in MPEA ML studies.^{39,84} Hyperparameter optimisation is also coupled with cross-validation to provide a more accurate assessment of the impact of hyperparameter changes on performance.^{33,43}

Table 3 summarises many recent studies where ML has been utilised to predict MPEA microstructural and mechanical properties, demonstrating the rapid adoption of ML in the field, in contrast to the lack of development in data availability, Table 2. As illustrated in Table 3 it is clear that the majority of ML models target alloy phase formation, because of the strong dependence of mechanical properties on microstructure. Of the alloy mechanical properties, hardness is the most commonly investigated in ML based MPEA studies, likely due to the ease of experimental measurement to validate model predictions and generate data.³

Phase formation prediction is a complex classification task with many possible phases that can form within

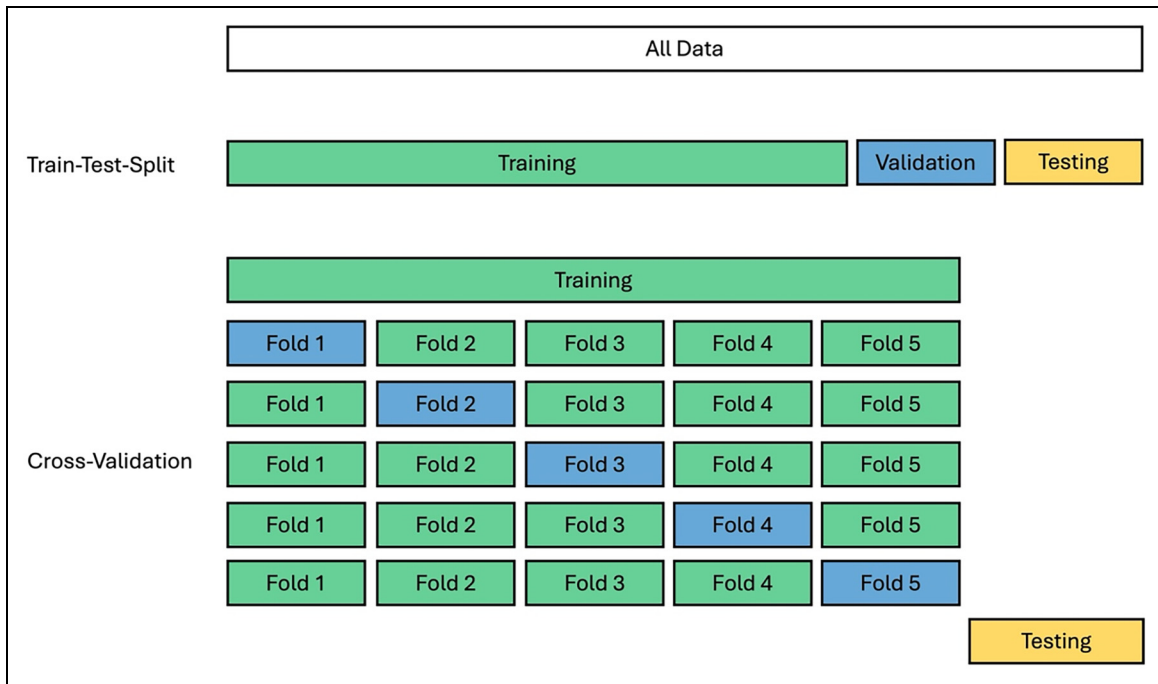


Figure 8. Comparison of the train-test-split and cross-validation model performance assessment methodologies. Highlights how cross-validation allows the model to be trained and validated multiple times before final testing.

MPEAs. Hence, many MPEA ML studies drastically condense the phase space in order to simplify the ML task and improve performance. For example, the MPEA database produced by Machaka et al.⁵⁶ contains 35 distinct labels for the phases of different compositions, which is simplified to 7 distinct labels. Furthermore, many MPEA ML studies reduce the classification problems down to binary and tertiary cases where the model is predicting whether or not solid solution formation occurs, or some form of solid solution, intermetallic or amorphous phase respectively.^{27,34,42,72,94} This reduction in the number of classes is helpful to obtain meaningful results from the ML, but can be exaggerated in the aim of maximising performance.

The studies in Table 3 mostly utilise standard ML algorithms with a small number of input features based on the intrinsic properties of the constituent elements. Furthermore, these studies typically utilise small databases, often with insufficient data to effectively train the more complex models, such as neural networks. In some cases these studies are lacking experimental validation of the model outputs, availability of data to enable future iterations, and sufficient consideration of the impact of manufacturing on MPEA design.

From machine learning to multi-principal element alloy design

Once the models have been successfully constructed, trained and their performance assessed, the next step is to use the model to predict on unseen data. To do this most MPEA studies generate a virtual candidate search space. This involves selecting the desired elements and

considering every possible combination in equimolar ratios, in a system of chosen size. For instance, Zhang et al.⁸⁷ considered 30 different elements in MPEA systems containing 4–6 elements in equimolar ratios. This produced a total of 763,686 unexplored alloy compositions for the ML to make predictions on the phase formation. This search space can be expanded to cover non-equimolar compositions, with each element typically being varied between 5 and 35 at.% with a compositional granularity of 1 at.%.^{84,86} In some MPEA studies, constraints are placed on the search space to further narrow it down. For example, in the study by Akhil et al.,³⁹ the search space was reduced by first applying the constraints, $\delta < 6.6\%$ and $-15 < H_{mix} < 5$ kJ/mol to aid in the search for HCP forming MPEAs. However, this is an unnecessary step as ML is a very fast process and if this constraint was suitable for mapping HCP formation, then it would be expected to be predicted accordingly by the ML model if trained correctly.

Alternatively, an inverse design methodology can be utilised.¹⁰⁴ In this case the final output property criteria is first specified as an input. The trained ML model then predicts feature configurations and hence, alloy compositions, that could satisfy the chosen property criteria.^{101,105} Yang et al.⁸⁴ utilised an inverse design methodology by projecting samples in the optimal property zone of the performance space and then deducing the features yielding this performance by inverse projection. Finally, the MPEA composition closest to the performance projection point by Euclidean distance was selected as the alloy composition to yield the high-performance property and was experimentally fabricated. From this inverse projection three MPEA compositions were downselected, synthesised and

Table 3. List of current publications on utilising supervised ML for MPEA design. In studies that compared multiple models, the best performing model or the one used to make predictions on unseen data is quoted in the table. Training data is experimentally determined unless otherwise stated. Data availability is the measure of how readily available the data for the published work is.

Author	Year	ML Target	ML Architecture	Training Dataset	Training Method	Performance Metrics	Num. Features	Data Availability
Islam et al. ⁴²	2018	Phase	Neural Network	118	4-Fold CV	Accuracy = 83.0%	5	Yes (Reference)
Wen et al. ⁸⁶	2019	Hardness	Support Vector Machine (RBF)	155	10-Fold CV	RMSE = 54.4	6	Yes
Chang et al. ⁹⁶	2019	Hardness	Neural Network	91	10-Fold CV	R ² = 0.94 MAE = 36 HV	3	No
Huang et al. ²⁷	2019	Phase	Artificial Neural Network	401	4-Fold CV	Accuracy = 74.3%	5	Yes
Choudhury et al. ⁹⁴	2020	Phase	Random Forest	119	TTS (70–30) 5-Fold CV	Accuracy = 91.6%	5	No
Kaufmann et al. ²⁸	2020	Phase	Random Forest	134 + 1664 DFT	5-Fold CV	Confidence = 75%	13	Yes
Roy et al. ⁹⁷	2020	Phase & Young's Modulus	Gradient Boosting	329 + 87	TTS (90–10)	MAE = 23.59 RMSE = 87.76	10	No
Pei et al. ⁵⁹	2020	Phase	Gaussian Process Classifier	1252	10-Fold CV	Accuracy = 93%	6	Yes
Zhang et al. ⁸⁷	2020	Phase	Support Vector Machine (RBF)	407	10-Fold CV	Accuracy = 75.3%	4	No
Akhil et al. ³⁹	2021	Phase	Extra Trees Classifier	111	Stratified 5-Fold CV	Accuracy = 91.6%	5	No
Bhandari et al. ⁴⁴	2021	Yield Strength	Random Forest	238	TTS (90–10) 10-Fold CV	Accuracy = 95%	4	No
Huang et al. ⁴⁰	2021	Hardness	Random Forest	106	10-Fold CV	MAE = 66.7HV RMSE = 86.2HV	13	Yes
Revi et al. ⁹⁸	2021	Elastic Constants	Random Forest	Materials Project	TTS (70–20–10)	R ² = 0.925 MAE = 13.14 RMSE = 19.61	5	No
Krishna et al. ⁷²	2021	Phase	Artificial Neural Network	636	TTS (75–25)	Accuracy = 80.5%	5	No
Machaka ⁸⁹	2021	Phase	Random Forest	896	TTS (75–25) 10-Fold CV	Accuracy = 97.5%	13	Yes
Han et al. ³⁴	2022	Phase	Extreme Gradient Boosting	867	TTS (80–20) 5-Fold CV	Accuracy > 85%	16	Yes
Li et al. ⁸⁸	2022	Hardness	Support Vector Machine (RBF)	205	10-Fold CV	RMSE = 47.9	4	Yes
Liu et al. ⁹⁹	2022	Primary Phase Fraction	Support Vector Machine	4 + 96 CALPHAD	TTS (80–20)	R ² = 0.916	–	No
Yang et al. ⁸⁴	2022	Hardness	Support Vector Machine (RBF)	370	TTS (80–20) 10-Fold CV	r = 0.94 RMSE = 75	5	Yes
Bakr et al. ⁴⁵	2022	Phase + Hardness	Neural Network + Ensemble of 3 Neural Networks	775 + 427	5 and 3-Fold CV	Accuracy = 93.4% R ² = 0.895 RMSE = 65.92	20	Yes
Chen et al. ⁴¹	2023	Hardness	Random Forest + Particle Swarm Optimisation	305	TTS (80–20)	R ² = 0.966 RMSE = 32.73	12	No
Sai et al. ¹⁰⁰	2023	Fatigue Life	Gradient Boost + Support Vector Machine	68 + 50	TTS (80–20)	Normalised RMSE = 0.07	9	No

(continued)

Table 3. Continued

Author	Year	ML Target	ML Architecture	Training Dataset	Training Method	Performance Metrics	Num. Features	Data Availability
Zhu et al. ³²	2023	Hardness	Deep Neural Network	324	4-Fold CV	MAE = 44.6	8	Yes (Reference)
Ren et al. ³³	2023	Hardness & Composition Optimisation	Random Forest + Support Vector Machine (RBF)	205	10-Fold CV	$R^2 = 0.9716$ RMSE = 39.2525	19	No
Jain et al. ⁴⁶	2023	Phase + Hardness	Extra Trees Classifier + Artificial Neural Network	1120 + 99	TTS (80–20)	Accuracy = 89.3% $R^2 = 0.95$ MAE = 34.91	21 and 11	No
Shen et al. ³⁸	2023	Phase & Hardness	Extreme Gradient Boosting	500+	3-Fold CV	Error Rate = 4.5% MAE = 26.49 RMSE = 52.66	7 and 8	No
Vazquez et al. ³⁵	2023	Phase Fraction	Deep Neural Network	229156 CALPHAD	TTS (70–10–20) 10-Fold CV	$R^2 = 0.95$	96	No
Wang et al. ¹⁰¹	2023	Yield Strength & Ultimate Tensile Strength	Convolutional Neural Network	501	Holdout Validation (Random Splits)	$R^2 = 0.866$ RMSE = 122	38	Yes
Berry et al. ¹⁰²	2024	Phase & Hardness	Random Forest	1360 + 370	5-Fold CV	Accuracy = 78.7% MAE = 66.1	10	Yes
Huang et al. ¹⁰³	2024	Ductility	Gradient Boosting	92	TTS (85–15)	Accuracy = 86% $R^2 = 0.85$	4	No

Definition of acronyms in table: Train-Test-Split (TTS), Cross-Validation (CV), Radial Basis Function kernel (RBF), CALculations of PHase Diagrams (CALPHAD), Density Functional Theory (DFT), Coefficient of Determination (R^2) and Root Mean Square Error (RMSE).

characterised. The best performing of these compositions exhibited a hardness 24.8% higher than the highest composition in the original dataset. Guo et al.¹⁰⁶ also apply inverse projection, but to optimise the composition of a preselected MPEA system, AlCrFeNiTi, to maximise hardness. From this inverse projection, four of the optimised composition alloys were experimentally synthesised. The highest hardness of these demonstrated a hardness 21.5% higher than any composition in the original dataset.

Debnath et al.¹⁰⁷ performed a comparative study of forward and inverse design methodologies with a case study on the ultimate tensile strength of refractory MPEAs. The authors observed that the inverse design methodology could identify a composition with the same constraints as the forward scheme, but produced a higher target property prediction. Additionally, the authors state that for the inverse design case the addition of new conditions and generation of new candidates is incredibly fast compared to the forward case, which needs to be regenerated with each iteration. However, inverse design is not without its disadvantages. Often numerous feasible solutions are presented by the inverse ML due to a greater number of variables than constraints and, minor variations in the desired property can produce significant changes in alloy composition prediction.¹⁰⁸

Experimental validation

As previously discussed, the success and development of ML within the MPEA field is intrinsically linked to the experimental exploration of the compositional space. However, a significant number of MPEA ML studies do not contain any form of experimental validation.^{28,39,44,45} The most common approach to experimental validation is to fabricate a small number of downselected samples, typically between 5 and 10, via vacuum arc or induction melting, or through powder metallurgy routes.¹⁰⁹

Once the samples have been fabricated, the standard approach is to investigate the microstructure and mechanical properties to validate the results of the ML predictions and discover new compositions with exceptional properties. The vast majority of MPEA ML studies perform this investigation via the application of scanning electron microscopy, in conjunction with energy dispersive x-ray spectroscopy, to determine the microstructure, compare the nominal to actual composition, and observe any elemental segregation. X-ray diffraction is also used to further crystallographically analyse the phase formation within the alloy. Mechanical property assessments, most commonly hardness, are performed using Vickers microhardness indentation.^{33,38,40,86,96} Examples of this experimental assessment are provided in Figure 9. Additional analysis techniques such as electron backscattered diffraction,³⁴ transmission electron microscopy and selected area diffraction⁹⁹ can enable greater insight into the crystallographic phase formation of the fabricated MPEAs, but this is typically above the resolution needed for ML model validation. However, there is a clear absence of extensive mechanical testing and post

processing of the alloys, with Wang et al.¹⁰¹ being one of the few to measure tensile properties of MPEAs following different post processing regimes predicted through ML.

The manufacturing methodology used to synthesise MPEAs significantly impacts alloy microstructure and phase formation, which governs the materials mechanical properties and overall performance.^{36,45,110,111} Post processing heat treatments can also further alter alloy microstructure and improve alloy mechanical performance.¹¹² For example, Otto et al.¹¹³ demonstrated that even the most exemplar single-phase solid solution MPEA system, CrMnFeCoNi, is not always single phase. After annealing at intermediate temperatures (<700°C), it will decay to form several secondary precipitates. Hence, incorporating the manufacturing and processing history into the ML task would assist the models to make accurate predictions irrespective of synthesis methodology.

For small scale trials of new MPEA compositions, arc melting is the dominant method for MPEA manufacture,¹¹⁴ especially for refractory MPEAs due to the high arc temperature.¹¹⁰ However, the formation of defects such as cold shuts,¹¹⁵ cracking and elemental segregation are common in arc-melted ingots.^{110,114} Thus, the microstructure and mechanical properties can vary significantly throughout the ingot in the as-cast state.¹¹⁴ Additive manufacturing processes offer faster cooling rates than conventional casting, resulting in a much finer grain structure, potentially enhancing mechanical properties.¹¹² However, the microstructures are more complex due to the different temperature histories and heating of subsequent layers.¹¹² In both cases, subsequent heat treatments can initiate phase transformations and improve mechanical properties. Homogenisation of MPEAs typically leads to higher strength than their as-cast counterparts¹¹¹ as well as mitigating or eliminating defects.¹¹²

Despite this, the majority of MPEA literature presents microstructure and mechanical properties in the as-cast condition,^{69,114} which is not indicative of industrial applications. Furthermore, the majority of MPEA ML studies don't appropriately account for processing conditions,^{45,116} select only as-cast data or omit any samples not in the as-cast state and manufactured via arc melting.^{33,40,84,86,87} The first barrier to the implementation of processing to MPEA ML studies is the lack of information available in publications, as it is not consistently reported within the literature.^{40,45} In addition, Machaka⁸⁹ reported that post processing heat-treatment features performed poorly in classification of phase formation tasks.

To incorporate the effect of manufacturing into the ML process Zhu et al.³² inserted the fabrication method into the input layer of a deep neural network for prediction of hardness by one-hot encoding. This is a process where categorical variables are mapped to integers so that they can be interpreted by the ML model. However, heat treated and laser remelted values of hardness were removed from the database in the initial data cleaning stage to eliminate the influence of process treatments. Two deep learning models were produced and the one incorporating

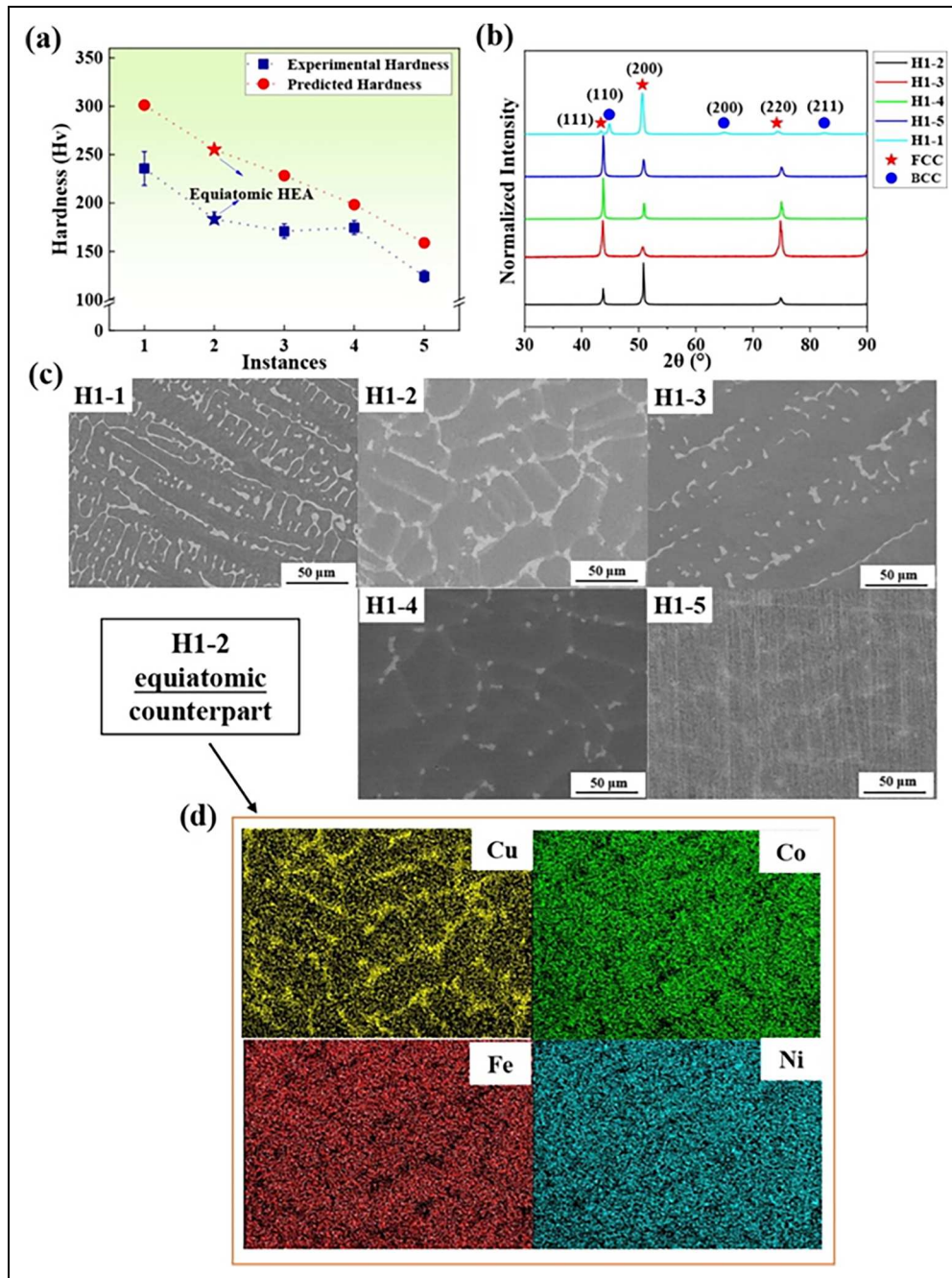


Figure 9. Experimental analysis performed on a series of 5 different compositions within the FeNiCuCo alloy system, downselected using a random forest model. a) Hardness. b) X-ray diffraction. c) Back scattered electron images. d) Energy dispersive x-ray spectroscopy maps of elemental distributions.⁴⁰ ©2024 reused with permission from Elsevier

manufacturing was found to perform better, highlighting the need to consider manufacturing conditions. Wang et al.¹⁰¹ employed a similar approach, but instead utilised processing conditions, such as annealing time and temperature, as input features due to their influence on mechanical properties. Bakr et al.⁴⁵ trained a series of neural networks, using nominal MPEA composition, manufacturing route and heat treatment temperature as the direct input features in the hardness model.

As discussed, when MPEA ML studies experimentally validate the ML model predictions they fabricate a small sample of compositions.³⁶ However, the combination of

large MPEA compositional space and capability of ML to rapidly make predictions across this unexplored space makes the connection of ML to High-Throughput Experiments (HTE) a logical advancement. HTE enable rapid synthesis and characterisations of MPEAs to validate ML predictions with enhanced efficiency.¹¹⁰ Thus, the application of HTE approaches to ML has the potential to further accelerate and reduce cost of MPEA design.¹¹⁷ Furthermore, the rapid generation of experimental data by implementation of HTEs can aid in constructing substantial MPEA databases for the development of future ML models.³⁶ Moorehead et al.¹¹⁸ utilised additive

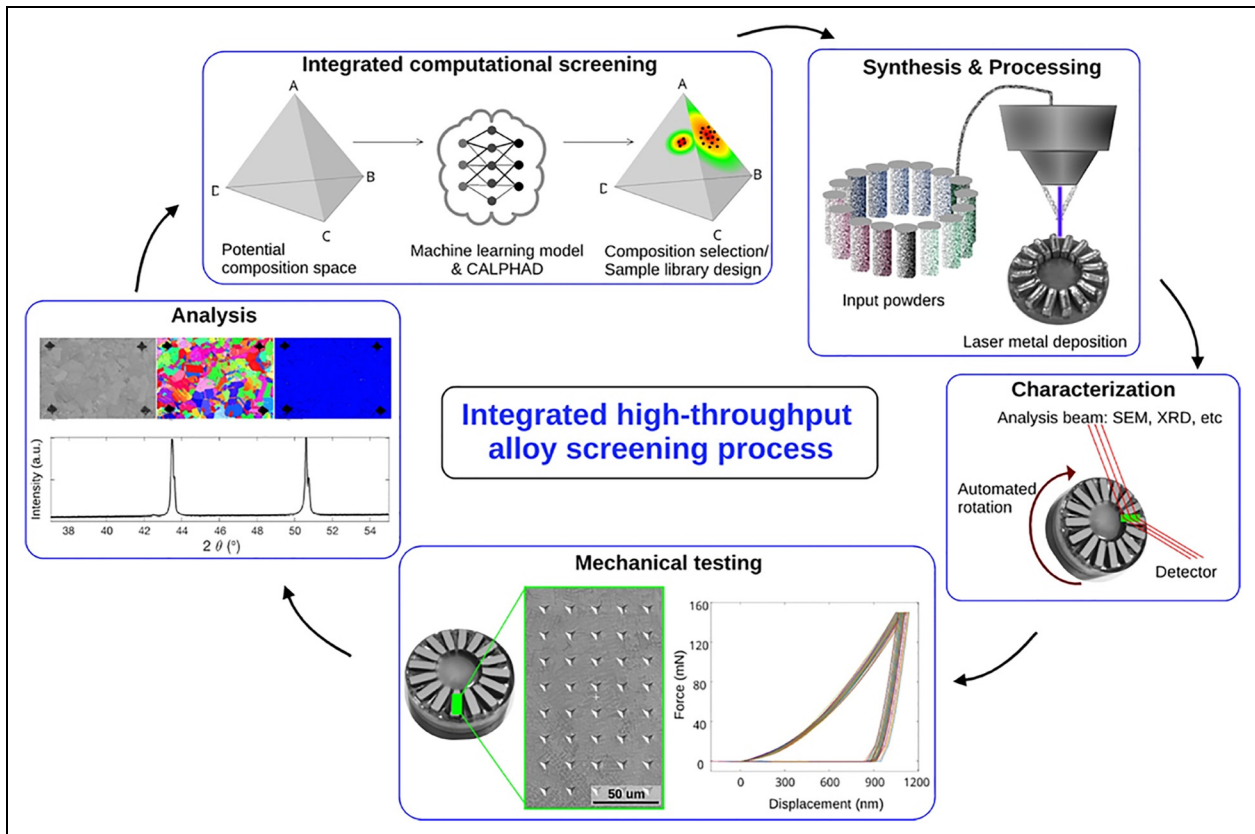


Figure 10. High-throughput experimental workflow from computation ML model predictions through alloy fabrication, characterisation, testing and analysis in an automated high-throughput manner, before closing the loop.¹¹⁹ ©2024 reused with permission from Elsevier

manufacturing in the form of directed energy deposition to construct arrays of different MPEA compositions onto a single build plate. 50 individual MPEA compositions are printed onto a single plate and x-ray diffraction, scanning electron microscopy and energy dispersive x-ray spectroscopy are all performed without removing the parts from the build plate. Moorehead et al.¹¹⁸ also produced an arc-melted button, indicative of those typically produced by MPEA ML studies for validation and found that additive manufacturing enabled a time saving of an order of magnitude compared to conventional casting. Vecchio et al.¹¹⁹ validated computational predictions of phase formation using CALPHAD, and hardness using an ML model with HTEs. MPEA samples were printed using directed energy deposition with a special hopper system allowing 16 different alloy compositions to be printed per build cycle. The MPEAs are built specifically to allow characterisation of microstructural and mechanical properties to also be performed in a high-throughput manner, Figure 10. In contrast, Liu et al.³⁶ applied HTE prior to ML to fabricate 138 MPEA compositions of the chosen alloy system through powder metallurgy. These compositions were then used to train an ML model to reveal the complete composition-hardness relationships across the compositional range of the chosen system. In addition, future expansion of HTE methodologies to include the implementation of autonomous or self-driving laboratories could further accelerate the discovery

and optimisation of materials,^{120–122} with MPEAs representing a suitable materials candidate.

Conclusions

Supervised ML has the potential to accelerate materials design and discovery within the MPEA space. The current direction of travel in the application of ML to the MPEA field is towards the complete automation of the materials design process. ML can be used to make large numbers of predictions on MPEA compositions to identify those with potentially advantageous mechanical and structural properties. These identified compositions can then be fabricated in a high-throughput manner within automated laboratories to enable model validation and assess their suitability to meet the design criteria. Regardless of the level of automation, domain knowledge of materials science is still integral to ensure the application of ML within the field is logical and incorporates the established understandings of decades of research. Additionally, there are many gaps identified within the field that need to be addressed in the interest of standardising and enhancing the use of supervised ML for MPEA design.

First and most critical, is the availability of MPEA data. Lack of available data is the most commonly cited reason for performance issues encountered in MPEA based ML studies. Hence, experimental MPEA data needs to be published and made open access to be utilised effectively in

future work within the field. This must include the method of fabrication and testing to improve the reliability of the data and enable it to be incorporated into ML models. The CALPHAD and DFT based data used in some studies to train models should also be made available to save on computationally expensive and time-consuming calculations. Furthermore, a wider, more general issue within the MPEA field is the need for a standardised naming convention for MPEAs to be established, making pre-existing data on alloy compositions easier to discern.

The relationship between supervised ML based alloy design methodologies and traditional alloy design methodologies needs to be considered. As discussed and shown in Table 3 of this report, a majority of ML MPEA studies focus on the optimisation of a single alloy property, either within a preselected system or across the whole MPEA space. In contrast, conventional alloy design processes look to balance many properties to achieve high performance while minimising cost and maintaining manufacturability. Hence, there is a clear need to move to multi output models.

Finally, the success and development of ML within the MPEA field is intrinsically linked to the experimental exploration of the compositional space. Many ML MPEA studies contain a workflow that portray data looping, where the experimental results of the study are fed back into the available data. However, the studies themselves end at the experimental assessments and publication. Thus, it is imperative that this data is not lost and is indeed recycled back into openly available databases to fuel further ML studies.

Acknowledgements

This work was supported by Oerlikon AM Europe GmbH (website: <https://www.oerlikon.com/am/en/>), Engineering and Physical Sciences Research Council UK [EP/S022635/1] (website: <https://www.ukri.org/councils/epsrc/>).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Engineering and Physical Sciences Research Council, Oerlikon AM Europe GmbH, (grant number EP/S022635/1).

ORCID iDs

Joshua Berry  <https://orcid.org/0000-0001-7291-2306>
Katerina A. Christofidou  <https://orcid.org/0000-0002-8064-5874>

References

1. Yeh JW, Chen SK, Lin SJ, et al. Nanostructured high-entropy alloys with multiple principal elements: novel

- alloy design concepts and outcomes. *Adv Eng Mater* 2004; 6: 299–303.
2. Cantor B, Chang ITH, Knight P, et al. Microstructural development in equiatomic multicomponent alloys. *Materials Science & Engineering A, Structural Materials: Properties, Microstructure and Processing* 2004; 375–377: 213–218.
3. Liu X, Zhang J and Pei Z. Machine learning for high-entropy alloys: progress, challenges and opportunities. *Progress in Materials Science* 2023; 131: 101018..
4. George EP, Raabe D and Ritchie RO. High-entropy alloys. *Nature Reviews Materials* 2019; 4: 515–534.
5. Miracle DB. Critical assessment 14: high entropy alloys and their development as structural materials. *Materials Science and Technology* 2015; 31: 1142–1147.
6. Miracle DB and Senkov ON. A critical review of high entropy alloys and related concepts. *Acta Materialia* 2017; 122: 448–511.
7. Miracle DB. High-Entropy alloys: a current evaluation of founding ideas and core effects and exploring “nonlinear alloys”. *JOM* 2017; 69: 2130–2136.
8. Pickering E and Jones N. High-Entropy alloys: a critical assessment of their founding principles and future prospects. *International Materials Reviews* 2016; 61: 183–202.
9. Molnar C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. 2nd Edition ed. 2022.
10. Pilania G. Machine learning in materials science: from explainable predictions to autonomous design. *Computational Materials Science* 2021; 193: 110360..
11. Zhang Y, Zhou YJ, Lin JP, et al. Solid-Solution phase formation rules for multi-component alloys. *Adv Eng Mater* 2008; 10: 534–538.
12. Yang X and Zhang Y. Prediction of high-entropy stabilized solid-solution in multi-component alloys. *Materials Chemistry and Physics* 2012; 132: 233–238.
13. Guo S, Ng C, Lu J, et al. Effect of valence electron concentration on stability of fcc or bcc phase in high entropy alloys. *Journal of Applied Physics* 2011; 109: 103505. doi: 10.1063/1.3587228
14. Guo S and Liu CT. Phase stability in high entropy alloys: formation of solid-solution phase or amorphous phase. *Progress in Natural Science: Materials International* 2011; 21: 433–446..
15. Guo S, Hu Q, Ng C, et al. More than entropy in high-entropy alloys: forming solid solutions or amorphous phase. *Intermetallics* 2013; 41: 96–103..
16. Wang YP, Li BS and Fu HZ. Solid solution or intermetallics in a high-entropy alloy. *Advanced Engineering Materials* 2009; 11: 641–644..
17. Ren M-x, Li B-s and Fu H-z. Formation condition of solid solution type high-entropy alloy. *Transactions of Nonferrous Metals Society of China* 2013; 23: 991–995.
18. Wang Z, Huang Y, Yang Y, et al. Atomic-size effect and solid solubility of multicomponent alloys. *Scripta Materialia* 2015; 94: 28–31.
19. Bracq G, Laurent-Brocq M, Perrière L, et al. The fcc solid solution stability in the Co-Cr-Fe-Mn-Ni multi-component system. *Acta Materialia* 2017; 128: 327–336.
20. Wu YD, Cai YH, Chen XH, et al. Phase composition and solid solution strengthening effect in TiZrNbMoV high-entropy alloys. *Materials & Design* 2015; 83: 651–660.
21. Liu WH, He JY, Huang HL, et al. Effects of Nb additions on the microstructure and mechanical property of CoCrFeNi high-entropy alloys. *Intermetallics* 2015; 60: –8.

22. Christofidou KA, McAuliffe TP, Mignanelli PM, et al. On the prediction and the formation of the sigma phase in CrMnCoFeNi_x high entropy alloys. *Journal of Alloys and Compounds* 2019; 770: 285–293.
23. Christofidou KA, Pickering EJ, Orsatti P, et al. On the influence of Mn on the phase stability of the CrMn_xFeCoNi high entropy alloys. *Intermetallics* 2018; 92: 84–92.
24. Rogal Ł, Szklarz Z, Bobrowski P, et al. Microstructure and mechanical properties of Al–Co–Cr–Fe–Ni base high entropy alloys obtained using powder metallurgy. *Metals and Materials International* 2019; 25: 930–945.
25. Morgan D and Jacobs R. Opportunities and challenges for machine learning in materials science. *Annual Review of Materials Research* 2020; 50: 71–103.
26. Hart GLW, Mueller T, Toher C, et al. Machine learning for alloys. *Nature Reviews Materials* 2021; 6: 730–755.
27. Huang W, Martin P and Zhuang HL. Machine-learning phase prediction of high-entropy alloys. *Acta Materialia* 2019; 169: 225–236.
28. Kaufmann K and Vecchio KS. Searching for high entropy alloys: a machine learning approach. *Acta Materialia* 2020; 198: 178–222.
29. Yan Y, Lu D and Wang K. Accelerated discovery of single-phase refractory high entropy alloys assisted by machine learning. *Computational Materials Science* 2021; 199: 110723.
30. Wang AY-T, Murdock RJ, Kauwe SK, et al. Machine learning for materials scientists: an introductory guide toward best practices. *Chemistry of Materials* 2020; 32: 4954–4965.
31. Maglogiannis IG. *Emerging artificial intelligence applications in computer engineering [electronic resource]: real word AI systems with applications in eHealth, HCI, information retrieval and pervasive technologies*. Amsterdam Washington, DC: IOS Press, 2007.
32. Zhu W, Huo W, Wang S, et al. Machine learning-based hardness prediction of high-entropy alloys for Laser additive manufacturing. *Jom* 2023; 75: 5537–5548.
33. Ren W, Zhang Y-F, Wang W-L, et al. Prediction and design of high hardness high entropy alloy through machine learning. *Materials & Design* 2023; 235: 112454. doi: 10.1016/j.matdes.2023.112454
34. Han Q, Lu Z, Zhao S, et al. Data-driven based phase constitution prediction in high entropy alloys. *Computational Materials Science* 2022; 215: 111774.
35. Vazquez G, Chakravarty S, Gurrola R, et al. A deep neural network regressor for phase constitution estimation in the high entropy alloy system al-co-cr-fe-mn-nb-ni. *npj Computational Materials* 2023; 9: 68.
36. Liu Y, Wang J, Xiao B, et al. Accelerated development of hard high-entropy alloys with data-driven high-throughput experiments. *Journal of Materials Informatics* 2022; 2: 3.
37. Zhang J, Cai C, Kim G, et al. Composition design of high-entropy alloys with deep sets learning. *npj Computational Materials* 2022; 8: 89. doi: 10.1038/s41524-022-00779-7
38. Shen L, Chen L, Huang J, et al. Predicting phases and hardness of high entropy alloys based on machine learning. *Intermetallics* 2023; 162: 108030.
39. Akhil B, Bajpai A, Gurao NP, et al. Designing hexagonal close packed high entropy alloys using machine learning. *Modelling Simul Mater Sci Eng* 2021; 29: 85005.
40. Huang X, Jin C, Zhang C, et al. Machine learning assisted modelling and design of solid solution hardened high entropy alloys. *Materials & Design* 2021; 211: 110177.
41. Chen C, Ma L, Zhang Y, et al. Accelerating the design of high-entropy alloys with high hardness by machine learning based on particle swarm optimization. *Intermetallics* 2023; 154: 107819.
42. Islam N, Huang W and Zhuang HL. Machine learning for phase selection in multi-principal element alloys. *Computational Materials Science* 2018; 150: 230–235.
43. Kaufmann K, Maryanovsky D, Mellor WM, et al. Discovery of high-entropy ceramics via machine learning. *npj Computational Materials* 2020; 6: 42.
44. Bhandari U, Rafi MR, Zhang C, et al. Yield strength prediction of high-entropy alloys using machine learning. *Materials Today Communications* 2021; 26: 101871. doi: 10.1016/j.mtcomm.2020.101871
45. Bakr M, Syarif J and Hashem IAT. Prediction of phase and hardness of HEAs based on constituent elements using machine learning models. *Materials Today Communications* 2022; 31: 103407. doi: 10.1016/j.mtcomm.2022.103407
46. Jain R, Lee U, Samal S, et al. Machine-learning-guided phase identification and hardness prediction of Al-Co-Cr-Fe-Mn-Nb-Ni-V containing high entropy alloys. *Journal of Alloys and Compounds* 2023; 956: 170193.
47. Zhang L, Chen H, Tao X, et al. Machine learning reveals the importance of the formation enthalpy and atom-size difference in forming phases of high entropy alloys. *Materials & Design* 2020; 193: 108835. doi: 10.1016/j.matdes.2020.108835
48. Debnath A, Krajewski AM, Sun H, et al. Generative deep learning as a tool for inverse design of high entropy refractory alloys. *Journal of Materials Informatics* 2021; 1: 3.
49. Jain A, Ong SP, Hautier G, et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials* 2013; 1: 011002. doi: 10.1063/1.4812323
50. Ye YF, Wang Q, Lu J, et al. High-entropy alloy: challenges and prospects. *Materials Today* 2016; 19: 349–362.
51. Gorsse S, Nguyen MH, Senkov ON, et al. Database on the mechanical properties of high entropy alloys and complex concentrated alloys. *Data Brief* 2018; 21: 2664–2678.
52. Gorsse S, Nguyen MH, Senkov ON, et al. Corrigendum to database on the mechanical properties of high entropy alloys and complex concentrated alloys, data in brief 21 (2018) 2664–2678. *Data in Brief* 2020; 32: 106216–106216.
53. Couzinié JP, Senkov ON, Miracle DB, et al. Comprehensive data compilation on the mechanical properties of refractory high-entropy alloys. *Data in Brief* 2018; 21: 1622–1641.
54. Balasubramanian A. Phases and Young's modulus dataset for High Entropy alloys. In: National Energy Technology Laboratory (NETL) P, PA, Morgantown, WV, and Albany, OR (United States). Energy Data eXchange; National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV (United States), (ed.). USDOE Office of Scientific and Technical Information2020.
55. Borg CKH, Frey C, Moh J, et al. Expanded dataset of mechanical properties and observed phases of multi-principal element alloys. *Scientific Data* 2020; 7: 30.
56. Machaka R, Motsi GT, Raganya LM, et al. Machine learning-based prediction of phases in high-entropy alloys: a data article. *Data in Brief* 2021; 38: 107346–107346.
57. Deter A, Oppenheimer S, Casey R, et al. Refractory high entropy alloy dataset with room temperature ductility screening. *Data in Brief* 2022; 45: 108582.

58. Chen S, Fan X, Steingrímsson B, et al. Fatigue dataset of high-entropy alloys. *Sci Data* 2022; 9: 381. 20220706.
59. Pei Z, Yin J, Hawk JA, et al. Machine-learning informed prediction of high-entropy solid solution formation: Beyond the Hume-Rothery rules. *npj computational materials* 2020; 6: 50. doi: 10.1038/s41524-020-0308-7
60. Beniwal D, Singh P, Gupta S, et al. Distilling physical origins of hardness in multi-principal element alloys directly from ensemble neural network models. *npj Computational Materials* 2022; 8: 153. doi: 10.1038/s41524-022-00842-3
61. Bundela AS and Rahul MR. Machine learning-enabled framework for the prediction of mechanical properties in new high entropy alloys. *Journal of Alloys and Compounds* 2022; 908: 164578. doi: 10.1016/j.jallcom.2022.164578
62. Ottomano F, De Felice G, Gusev VV, et al. Not as simple as we thought: a rigorous examination of data aggregation in materials informatics. *Digital Discovery* 2024; 3: 337–346.
63. Lee SY, Byeon S, Kim HS, et al. Deep learning-based phase prediction of high-entropy alloys: Optimization, generation, and explanation. *Materials & Design* 2021; 197: 109260. doi: 10.1016/j.matdes.2020.109260
64. Gilligan LPJ, Cobelli M, Taufour V, et al. A rule-free workflow for the automated generation of databases from scientific literature. *npj Computational Materials* 2023; 9: 222. doi: 10.1038/s41524-023-01171-9
65. Dagdelen J, Dunn A, Lee S, et al. Structured information extraction from scientific text with large language models. *Nature Communications* 2024; 15: 1418. doi: 10.1038/s41467-024-45563-x
66. Risal S, Zhu W, Guillen P, et al. Improving phase prediction accuracy for high entropy alloys with Machine learning. *Computational Materials Science* 2021; 192: 110389. doi: 10.1016/j.commatsci.2021.110389
67. Elreedy D and Atiya AF. A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance. *Information Sciences* 2019; 505: 32–64.
68. Bansal A, Kumar P, Yadav S, et al. Accelerated design of high entropy alloys by integrating high throughput calculation and machine learning. *Journal of Alloys and Compounds* 2023; 960: 170543. doi: 10.1016/j.jallcom.2023.170543
69. Singh S, Katiyar NK, Goel S, et al. Phase prediction and experimental realisation of a new high entropy alloy using machine learning. *Sci Rep* 2023; 13: 4811. 20230323.
70. Hareharen K, Panneerselvam T and Raj Mohan R. Improving the performance of machine learning model predicting phase and crystal structure of high entropy alloys by the synthetic minority oversampling technique. *Journal of Alloys and Compounds* 2024; 991: 174494. doi: 10.1016/j.jallcom.2024.174494
71. Hastie T, Friedman JH and Tibshirani R. *The elements of statistical learning : data mining, inference, and prediction*. 2nd ed. New York: Springer, 2009.
72. Krishna YV, Jaiswal UK and R RM. Machine learning approach to predict new multiphase high entropy alloys. *Scripta Materialia* 2021; 197: 113804.
73. Kube SA, Sohn S, Uhl D, et al. Phase selection motifs in high entropy alloys revealed through combinatorial methods: large atomic size difference favors BCC over FCC. *Acta Materialia* 2019; 166: 677–686.
74. Tian F, Varga LK, Chen N, et al. Empirical design of single phase high-entropy alloys with high hardness. *Intermetallics* 2015; 58: –6.
75. Fang S, Xiao X, Xia L, et al. Relationship between the widths of supercooled liquid regions and bond parameters of Mg-based bulk metallic glasses. *Journal of non-Crystalline Solids* 2003; 321: 120–125.
76. Mizutani U and Sato H. The physics of the Hume-Rothery electron concentration rule. *Crystals (Basel)* 2017; 7: 9.
77. Calvo-Dahlborg M, Mehraban S, Lavery NP, et al. Prediction of phase, hardness and density of high entropy alloys based on their electronic structure and average radius. *Journal of Alloys and Compounds* 2021; 865: 158799.
78. Takeuchi A and Inoue A. Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element. *Mater Trans* 2005; 46: 2817–2829.
79. Rickman JM, Balasubramanian G, Marvel CJ, et al. Machine learning strategies for high-entropy alloys. *Journal of Applied Physics* 2020; 128: 221101. doi: 10.1063/5.0030367
80. Callister WD and Rethwisch DG. *Materials science and engineering. Ninth edition, SI version*. ed. Hoboken, New Jersey, Hoboken, NJ: Wiley, 2015.
81. Prieto E, Vaz-Romero A, Gonzalez-Julian J, et al. Novel high entropy alloys as binder in cermets: from design to sintering. *International Journal of Refractory Metals & Hard Materials* 2021; 99: 105592.
82. Gwalani B, Choudhuri D, Liu K, et al. Interplay between single phase solid solution strengthening and multi-phase strengthening in the same high entropy alloy. *Materials Science & Engineering A, Structural Materials : Properties, Microstructure and Processing* 2020; 771: 138620.
83. George EP, Curtin WA and Tasan CC. High entropy alloys: a focused review of mechanical properties and deformation mechanisms. *Acta Materialia* 2020; 188: 435–474.
84. Yang C, Ren C, Jia Y, et al. A machine learning-based alloy design system to facilitate the rational design of high entropy alloys with enhanced hardness. *Acta Materialia* 2022; 222: 117431.
85. Zeng Y, Man M, Bai K, et al. Revealing high-fidelity phase selection rules for high entropy alloys: a combined CALPHAD and machine learning study. *Materials & Design* 2021; 202: 109532.
86. Wen C, Zhang Y, Wang C, et al. Machine learning assisted design of high entropy alloys with desired property. *Acta Materialia* 2019; 170: 109–117.
87. Zhang Y, Wen C, Wang C, et al. Phase prediction in high entropy alloys with a rational selection of materials descriptors and machine learning models. *Acta Materialia* 2020; 185: 528–539.
88. Li S, Li S, Liu D, et al. Hardness prediction of high entropy alloys with machine learning and material descriptors selection by improved genetic algorithm. *Computational Materials Science* 2022; 205: 111185.
89. Machaka R. Machine learning-based prediction of phases in high-entropy alloys. *Computational Materials Science* 2021; 188: 110244.
90. James G, Witten D, Hastie T, et al. *An introduction to statistical learning : with applications in R*. New York: New York: Springer, 2013.
91. Campbell MJ. *Statistics at square one*. Twelfth edition. Campbell Michael J. (ed.) Hoboken, NJ: Wiley Blackwell, 2021, pp.165–169.

92. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med* 2018; 18: 91–93. 20180807.
93. Papageorgiou SN. On correlation coefficients and their interpretation. *Journal of Orthodontics* 2022; 49: 359–361.
94. Choudhury A, Konnur T, Chattopadhyay PP, et al. Structure prediction of multi-principal element alloys using ensemble learning. *Engineering Computations* 2020; 37: 1003–1022.
95. Dalianis H. Evaluation metrics and evaluation. In: Dalianis H (eds) *Clinical text mining: secondary use of electronic patient records*. Cham: Springer International Publishing, 2018, pp.45–53.
96. Chang Y-J, Jui C-Y, Lee W-J, et al. Prediction of the composition and hardness of high-entropy alloys by machine learning. *JOM* 2019; 71: 3433–3442.
97. Roy A, Babuska T, Krick B, et al. Machine learned feature identification for predicting phase and Young's modulus of low-, medium- and high-entropy alloys. *Scripta Materialia* 2020; 185: 152–158.
98. Revi V, Kasodariya S, Talapatra A, et al. Machine learning elastic constants of multi-component alloys. *Computational Materials Science* 2021; 198: 110671. doi: 10.1016/j.commatsci.2021.110671
99. Liu F, Xiao X, Huang L, et al. Design of NiCoCrAl eutectic high entropy alloys by combining machine learning with CALPHAD method. *Materials Today Communications* 2022; 30: 103172.
100. Sai NJ, Rathore P and Chauhan A. Machine learning-based predictions of fatigue life for multi-principal element alloys. *Scripta Materialia* 2023; 226. DOI: 10.1016/j.scriptamat.2022.115214
101. Wang J, Kwon H, Kim HS, et al. A neural network model for high entropy alloy design. *npj Computational Materials* 2023; 9: 60.
102. Berry J, Snell R, Anderson M, et al. Design and selection of high entropy alloys for hardmetal matrix applications using a coupled machine learning and calculation of phase diagrams methodology. *Advanced Engineering Materials* 2024; 26: 2302064. doi: 10.1002/adem.202302064
103. Huang X, Zheng L, Xu H, et al. Predicting and understanding the ductility of BCC high entropy alloys via knowledge-integrated machine learning. *Materials & Design* 2024; 239: 112797. doi: 10.1016/j.matdes.2024.112797
104. Lee C-Y, Jui C-Y, Yeh A-C, et al. Inverse design of high entropy alloys using a deep interpretable scheme for materials attribution analysis. *Journal of Alloys and Compounds* 2024; 976: 173144. doi: 10.1016/j.jallcom.2023.173144
105. Zeng Y, Man M, Ng CK, et al. Machine learning-based inverse design for single-phase high entropy alloys. *APL Materials* 2022; 10: 112929. doi: 10.1063/5.0109491
106. Guo Q, Pan Y, Hou H, et al. Predicting the hardness of high-entropy alloys based on compositions. *International Journal of Refractory Metals and Hard Materials* 2023; 112. DOI: 10.1016/j.ijrmhm.2023.106116
107. Debnath A, Raman L, Li W, et al. Comparing forward and inverse design paradigms: a case study on refractory high-entropy alloys. *Journal of Materials Research* 2023; 38: 4107–4117.
108. Lee J, Park D, Lee M, et al. Machine learning-based inverse design methods considering data characteristics and design space size in materials design and manufacturing: a review. *Mater Horiz* 2023; 10: 5436–5456. 20231127.
109. Kaushik N, Meena A and Mali HS. High entropy alloy synthesis, characterisation, manufacturing & potential applications: a review. *Materials and Manufacturing Processes* 2021; 37: 1085–1109.
110. Zhang Y and Xing Q. High entropy alloys: manufacturing routes. In: Caballero FG (eds) *Encyclopedia of materials: metals and alloys*. Oxford: Elsevier, 2022, pp.327–338.
111. Alshataif YA, Sivasankaran S, Al-Mufadi FA, et al. Manufacturing methods, microstructural and mechanical properties evolutions of high-entropy alloys: a review. *Metals and Materials International* 2019; 26: 1099–1133.
112. Zhang W, Chabok A, Kooi BJ, et al. Additive manufactured high entropy alloys: A review of the microstructure and properties. *Materials & Design* 2022; 220: 110875. doi: 10.1016/j.matdes.2022.110875
113. Otto F, Dlouhý A, Pradeep KG, et al. Decomposition of the single-phase high-entropy alloy CrMnFeCoNi after prolonged anneals at intermediate temperatures. *Acta Materialia* 2016; 112: 40–52.
114. Jablonski PD, Licavoli JJ, Gao MC, et al. Manufacturing of high entropy alloys. *Jom* 2015; 67: 2278–2287.
115. Nagase T, Mizuuchi K and Nakano T. Solidification microstructures of the ingots obtained by arc melting and cold crucible levitation melting in TiNbTaZr Medium-entropy alloy and TiNbTaZrX (X = V, Mo, W) high-entropy alloys. *Entropy (Basel)* 2019; 21: 20190510.
116. Wang Z, Li L, Chen Z, et al. A new route to achieve high strength and high ductility compositions in Cr-Co-Ni-based medium-entropy alloys: a predictive model connecting theoretical calculations and experimental measurements. *Journal of Alloys and Compounds* 2023; 959: 170555.
117. Huang EW, Lee W-J, Singh SS, et al. Machine-learning and high-throughput studies for high-entropy materials. *Materials Science and Engineering: R: Reports* 2022; 147: 100645. doi: 10.1016/j.mser.2021.100645
118. Moorehead M, Bertsch K, Niezgoda M, et al. High-throughput synthesis of Mo-Nb-Ta-W high-entropy alloys via additive manufacturing. *Materials & Design* 2020; 187: 108358. doi: 10.1016/j.matdes.2019.108358
119. Vecchio KS, Dippo OF, Kaufmann KR, et al. High-throughput rapid experimental alloy development (HT-READ). *Acta Materialia* 2021; 221. doi: 10.1016/j.actamat.2021.117352
120. Baird SG and Sparks TD. What is a minimal working example for a self-driving laboratory? *Matter* 2022; 5: 4170–4178.
121. MacLeod BP, Parlange FGL, Rupnow CC, et al. A self-driving laboratory advances the Pareto front for material properties. *Nat Commun* 2022; 13: 995. 20220222.
122. Szymanski NJ, Rendy B, Fei Y, et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* 2023; 624: 86–91. 20231129.