

Article

Machine Learning Phase Prediction of Light-Weight High-Entropy Alloys Containing Aluminum, Magnesium, and Lithium

Shenglong Li ^{1,2,3}, Rui Liu ^{1,2,3}, Hongwei Yan ^{1,2,3,*}, Zhihui Li ^{1,3,*}, Yanan Li ^{1,2,3}, Xiwu Li ^{1,2,3}, Yongan Zhang ^{1,2,3} and Baiqing Xiong ^{1,3}

¹ State Key Laboratory of Nonferrous Metals and Processes, GRINM Group Co., Ltd., Beijing 100088, China

² GRIMAT Engineering Institute Co., Ltd., Beijing 101407, China

³ General Research Institute for Nonferrous Metals, Beijing 100088, China

* Correspondence: yanhongwei@grinm.com (H.Y.); lzh@grinm.com (Z.L.)

Abstract: With the development of society, there is an increasingly urgent demand for light-weight, high-strength, and high-temperature-resistant structural materials. High-entropy alloys (HEAs) owe much of their unusual properties to the selection among three phases: solid solution (SS), intermetallic compound (IM), and mixed SS and IM (SS and IM). Therefore, accurate phase prediction is crucial for guiding the selection of element combinations to form HEAs with desired properties. Light high-entropy alloys (LHEAs), as a significant branch of HEAs, exhibit excellent performance in terms of specific strength. In this study, we employ a machine learning (ML) method to realize the design of light-weight high-entropy alloys based on solid solutions. We determined the Gradient Boosting Classifier model as the best machine learning model through a two-step feature and model selection, in which its accuracy and F1_Score achieve 0.9166 and 0.8923. According to the predicted results, we obtained $\text{Al}_{28}\text{Li}_{35}\text{Mg}_{15}\text{Zn}_{10}\text{Cu}_{12}$ LHEAs, which are mainly composed of 90% solid solution. This alloy accords with the prediction results of machine learning. But it is made up of a two-phase solid solution. In order to obtain a light-weight high-entropy alloy dominated by a single solid solution, we designed $\text{Al}_{24}\text{Li}_{15}\text{Mg}_{26}\text{Zn}_9\text{Cu}_{26}$ LHEAs on the basis of machine learning prediction results accompanied by expert experience. Its main structure includes a single-phase solid solution. Our work provides an alternative approach to the computational design of HEAs and provides a direction for future exploration of light-weight high-entropy alloys.

Keywords: light-weight high-entropy alloys; machine learning; alloy design; structure prediction



Citation: Li, S.; Liu, R.; Yan, H.; Li, Z.; Li, Y.; Li, X.; Zhang, Y.; Xiong, B.

Machine Learning Phase Prediction of Light-Weight High-Entropy Alloys Containing Aluminum, Magnesium, and Lithium. *Metals* **2024**, *14*, 400. <https://doi.org/10.3390/met14040400>

Academic Editors: Jiro Kitagawa, João Manuel R. S. Tavares and Babak Shalchi Amirkhiz

Received: 21 February 2024

Revised: 12 March 2024

Accepted: 26 March 2024

Published: 29 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The pursuit of high-strength light-weight structural components has always been an advanced research focus in the aerospace industry, especially with the blooming development of new-generation high-demanding warcraft, such as supersonic aircraft, high-demanding warcraft, space vehicles, etc. While traditional high-strength aluminum alloys can meet the corresponding performance requirements, their strength and elastic modulus drop rapidly at high temperatures, rendering them to service in extreme application environments [1]. Magnesium and titanium alloys are also limited by their low strength and they are cost-prohibitive, which are imposed restrictions as structural components in the aerospace industry. Additionally, the development of traditional alloy systems is nearly saturated, which makes a difficult increment create new alloy systems owing to the constraints of the design principles for traditional alloys.

In the early 21st century, Yeh and other researchers introduced the concept of high-entropy alloys (HEAs), defining them as alloys containing at least five principal elements, with the atomic percentage of each major element falling between 5% and 35% [2–4]. Multi-principal element high-entropy alloys, after solidification, tend to form structurally

simple solid solutions, contrary to the formation of numerous intermetallic compounds, as traditional concepts suggest. Compared with traditional alloys, high-entropy alloys have excellent properties, such as high strength [5–8], high hardness [9–11], high corrosion resistance [12–14] and high-temperature resistance [15–17]. Since it was proposed in 2004, HEAs have attracted wide interest not only because of their performance advantages, but also because they open up a vast new field in alloy design.

As a significant branch of HEAs, LHEA alloys possess outstanding specific strength advantages compared to traditional light-weight alloys and show promising application potential in future extreme environments and military fields. Nevertheless, in the early research on Light high-entropy alloys (LHEAs), due to the lack of guiding empirical criteria, most publicly reported that light-weight high-entropy alloys have complex structures and poor plasticity [18–20]. Whether HEA exhibits excellent properties depends on which phase HEA is used in solid solutions (SS), intermetallic compounds (IM), and mixtures of SS and IM (SS + IM). It is precisely because of the existence of the SS phase that HEAs have the above excellent performance. In contrast, the presence of a brittle IM phase reduces ductility [21]. Therefore, accurate prediction of the results of a given combination of components is critical to the development and application of new HEAs. Moreover, due to compositional characteristics, the design of HEAs faces a high-dimensional and complex search space, which makes it difficult to rapidly and accurately design reasonable alloy compositions by traditional empirical trial-and-error methods. The machine learning method exhibits the capability to extract new information from existing experimental data, explore complex hidden relationships among various parameters, establish accurate predictive models, and fully leverage the role of experimental data. Clearly, this emerging method can effectively improve the speed of developing new alloy compositions. Li et al. [22] established a dataset of 322 as-cast alloy compositions and accurately differentiated FCC (face-centered cubic), BCC (body-centered cubic), and NSP (not forming single-phase solid solution) phases using an optimized support vector machine model, achieving a model accuracy of up to 90%. Nong et al. [23] studied the stability of cubic structures from the perspective of atomic size differences, and established a backpropagation artificial neural network based on the calculated solid solution physical parameters and structural characteristics of cast HEAs to predict alloy structures. The predicted values of the training set and the test set are in good agreement with the experimental values, and the correlation coefficients R^2 are 0.988 and 0.824, respectively, indicating that the performance of BP-ANN has accurately predicted the structure of cast high-entropy alloys. Huang et al. [24] used three different machine learning algorithms, K-nearest neighbors (KNN), support vector machines (SVM), and artificial neural networks (ANN), to cross-validate an experimental dataset which found that the trained ANN model performed the best among the three ML algorithms, achieving an accuracy of up to 94.3%. Despite extensive research on machine learning-assisted material design, there are no reported studies on the design of ultra-light-weight HEAs (AlLiMg system).

In this work, as shown in Figure 1, we combine a data-driven approach with knowledge in the field of HEAs to achieve the designs of LHEAs, with solid solution phases dominating. The experimental validation was conducted for the microstructures of alloys to demonstrate the reliability of the material design approach assisted by machine learning, providing guidance for the development of low-density LHEAs.

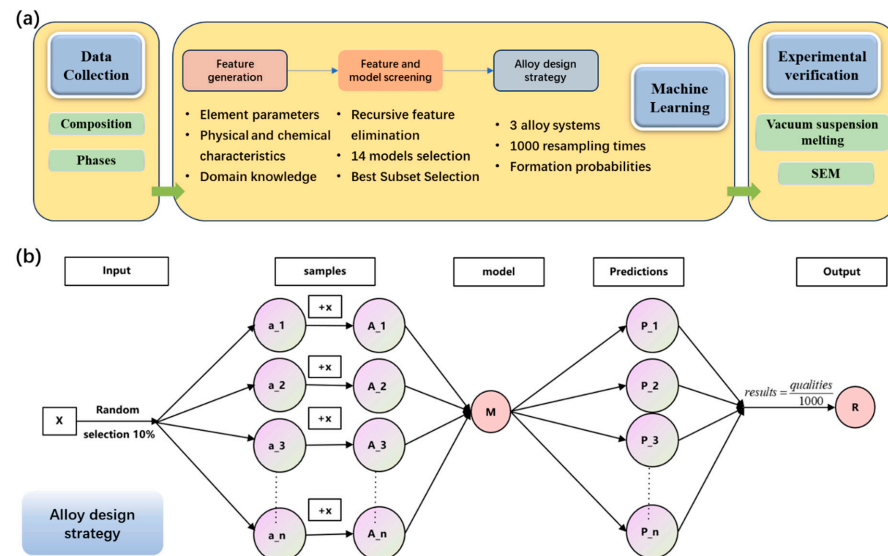


Figure 1. Rational design approach of new light-weight high entropy alloy. (a) Machine learning process; (b) alloy design strategy.

2. Machine Learning Design Strategy

2.1. Data Collection and Analysis

The development of HEAs has a history of less than two decades and has not established comprehensive databases (such as MatWeb [25] and other standard databases). Therefore, the dataset of LHEAs was built by data collection from the literature. Because research on LHEAs is still in its preliminary stage, there is no universally agreed-upon definition [26–28]. This study summarizes that the reported densities of LHEAs generally range from 2 to 6 g/cm³. After removing redundant and incomplete data, we collected 97 data with densities below 6 g/cm³, including atomic percentages, mass percentages, types of phases, and processing methods. The elemental distribution of all alloys is illustrated in Figure 2 where Al, Li, Mg, Zn, Cu, and Ti are the predominant elements in the dataset.

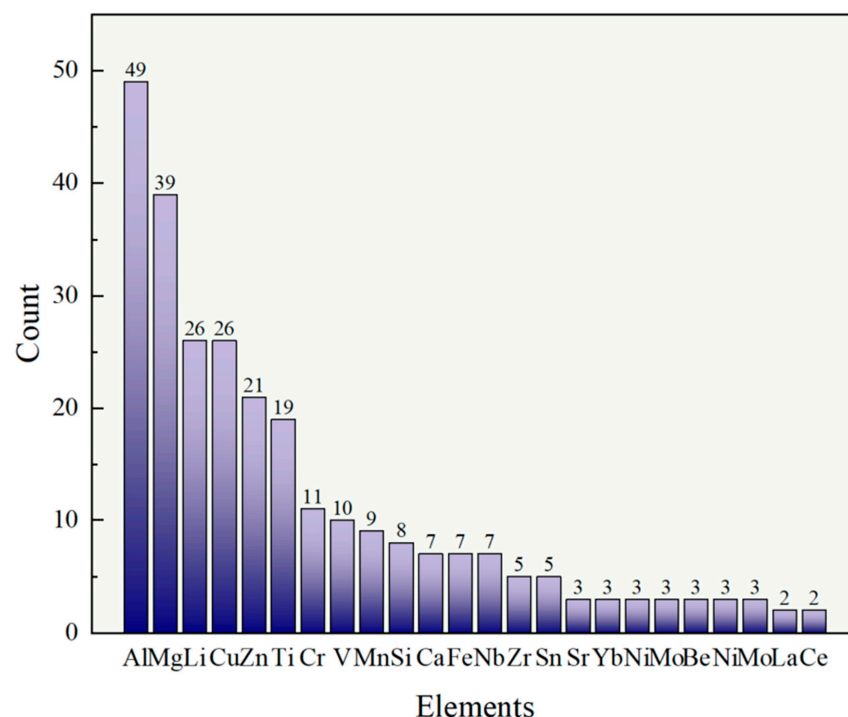


Figure 2. Histogram of phase element distribution.

Currently, most LHEAs display complex microstructure, especially in the AlLiMg system. Therefore, we established a set of rules in the process of data classification: if the proportion of SS in the alloy is above 90%, it is classified as a single-phase solid solution; if the proportion of SS is between 50% and 90%, it is identified as a combination of solid solution and intermetallic compound phases; if the proportion of SS is less than 50%, it is considered a structure of intermetallic compounds. By applying these rules, the dataset is divided into 22 data of SS, 39 data of SS and IM, and 36 data of IM. These three types of phase structures are represented by 0, 1, and 2 as classification labels in machine learning models. The dataset includes ternary alloys, quaternary alloys, quinary alloys, senary alloys, and septenary alloys, indicating the broad scope of the data. The quantities of different elements and the numbers of different phases in the dataset are shown in Figure 2 and Table 1.

Table 1. Components and phase distribution of different alloys (SS: solid solution, IM: intermetallic).

Alloy	SS	SS + IM	IM	Total
Ternary	0	1	5	6
Quaternary	6	2	4	12
Five elements	10	33	21	64
Six elements	3	0	3	6
Seven elements	3	3	3	9

2.2. Screening Method

Researchers have revealed that various factors can have different impacts on the phase formation of HEAs, such as mixing enthalpy (ΔH_{mix}), mixing entropy (ΔS_{mix}), atomic radius difference (δ), valence electron concentration (VEC), electronegativity (χ), Ω parameter, etc. And they have summarized general empirical criteria for phase formation in HEAs by extensive exploration [29]: $\delta < 6.5\%$, $-15 \text{ kJ/mol} < \Delta H_{mix} < 5 \text{ kJ/mol}$, $12 \text{ J/(K}\cdot\text{mol)} < \Delta S_{mix} < 17.5 \text{ J/(K}\cdot\text{mol)}$, $\Omega \geq 1.1$. However, it has been observed that these phase formation rules may not be effective for LHEAs containing high concentrations of light-weight elements such as Al, Li, Mg, etc. In multicomponent alloys with a significant amount of low-density elements, configurational entropy does not seem to be the primary factor controlling phase selection, and low-density multicomponent alloys are not prone to forming simple solid solutions. Yang et al. [20] suggested that it is challenging to form disordered solid solutions in low-density AlLiMg-LHEAs, despite the existence of configurational entropy effects. This difficulty may be attributed to the differences in bonding characteristics between these elements or the lack of d orbitals in low-density elements hindering higher-order hybridization. Therefore, in addition to the features mentioned above, it is essential to incorporate as many potential influencing factors as possible that may impact the model to better identify useful features.

Alloys have various parameters and properties that can be used as input features, including elemental metrics (atomic number, period, group, orbital electron number, unfilled orbital electron number, effective nuclear charge, etc.), physical properties (relative molecular mass, van der Waals radius, atomic radius, ionic radius, covalent radius, density, melting point, boiling point, weight, etc.), and electrochemical properties (electronegativity, valence electron number, ionization energy, etc.) [30]. The feature space constructed in this study contains a wealth of physical and chemistry information, while also involving knowledge from the HEAs field. Finally, 81 feature candidates were generated as inputs of the ML model. Among them, 17 HEA features were referenced from descriptors constructed by Wen [31], while other features were generated using ElementProperty, YangSolidSolution, Miedema, and ValenceOrbital toolkits, calculating the mean and average deviation of alloy element information. In order to accelerate the computational speed of the model and

enhance accuracy, we conducted dimensionality reduction on the features [32], and the generated features were normalized using the following formula:

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

where x represents the original feature, x_{norm} represents the normalized feature, and $\max(x)$ and $\min(x)$ are the maximum and minimum values of feature x , respectively.

2.3. Feature and Model Selection

After completing data preprocessing, it is crucial for training machine learning algorithms and models to select meaningful features, as the quality of features significantly influences the model's performance. Feature filtering is eliminating redundant and unrelated features from the original feature space, retaining only the useful ones [33]. The right choice of machine learning model can significantly reduce model complexity and save computational time. In our work, we designed a two-step process for feature and model selection.

The first step utilized Recursive Feature Elimination (RFE) [34], which selects features based on the importance obtained from the `coef_` or `feature_importances_` attribute returned by the learner. The least important features are iteratively removed from the feature space during model training until the desired number of features is reached. Here, we chose the Gradient Boosting Classifier model to achieve this process.

The second step is model selection [35]. In order to avoid the Gradient Boosting Classifier model is not the best optimal choice, we trained 14 different classification models, including Logistic Regression (LR), Decision Tree Classifier (DT), Support Vector Classifier (SVC), Random Forest Classifier (RF), and Gaussian NB (GNB), K-nearest Neighbors Classifier (KNN), AdaBoost Classifier (AB), Gradient Boosting Classifier (GB), XGB Classifier XGB, LGBM Classifier (LGB), Cat Boost Classifier (CB), MLP Classifier (MLP), SGD Classifier (SGD), and Gaussian Process Classifier (GP). Finally, we selected the three models with the best performance for the next step.

The third step is carried out according to the Best Subset Selection (BSS) [36] method for model and feature screening. This method involves exhaustively evaluating multiple models and features by traversing all possible feature combinations, resulting in the best model and corresponding optimal feature subset. Models and features selected through this method are guaranteed to be the best because they explore all possibilities.

During the model and feature selection process, evaluation metrics and methods are necessary to assess the model's performance. We employed ten-fold cross-validation to split the dataset, a method that provides a more reasonable and accurate evaluation, especially when dealing with small datasets [37]. The performance of the model is determined by accuracy and F1_Score. Accuracy is the most intuitive evaluation metric, representing the proportion of correctly predicted samples to the total number of samples. The calculation is as follows:

$$accuracy(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=1}^{n_{samples}-1} l(\hat{y}_i = y_i) \quad (2)$$

where y is the true value of the sample, $\hat{y}$ is the predicted value, $n_{samples}$ is the number of samples, and $l(x)$ is the indicator function.

The F1_Score is a metric that combines both *precision* and *recall*, providing a balanced assessment of a model's performance, especially in situations where there is an imbalance between classes. *Precision* and *recall* are defined as follows:

$$precision = \frac{true\ positive}{true\ positive + false\ positive} \quad (3)$$

$$recall = \frac{true\ positive}{true\ positive + false\ negative} \quad (4)$$

In the context provided, *true positive* represents correct results, *false positive* indicates unexpected results, and *false negative* signifies missing results. The calculation method for the F1_Score involves taking the harmonic mean of *precision* and *recall*:

$$F1_score = (1 + \beta^2) \frac{precision \times recall}{precision + recall} \quad (5)$$

where the choice of the β parameter determines the relative emphasis on *precision* and *recall*.

2.4. Experimental Methods

The experimental alloy ingots were prepared through vacuum levitation melting. Considering the volatility of Li and Mg elements, we prepared AlLi₂₀ and MgLi₂₀ in advance, instead of pure metals. Smelting materials are small particles or small pieces. The purity of all Al, Li, Mg, Zn, and Cu raw materials is 99.99%. Due to differences in melting points and densities among these elements, the melting process was divided into multiple stages that added materials from high to low according to the melting point. Each alloy ingot was melted five times. It is worth noting that the alloy ingot needs to be weighed after each melting and the volatile elements should be replenished at the next melting to ensure that the composition is controlled within a reasonable range. In order to analyze the microstructure of the alloys, the field emission scanning electron microscope (SEM, JSM-7900F, Tokyo, Japan) equipped with an Octane Elect Super energy dispersion spectrometer (EDS) was used, which could further distinguish different elements and calculate the relative content of elements, so as to obtain the content and type of elements in different phase regions. But, EDS cannot analyze elements with relative atomic mass less than 7, such as the Li element of the LHEAs in this study.

3. Results and Discussion

3.1. Construction of the Model and Analysis of the Design Results

The results of the RFE process are illustrated in Figure 3a, which shows the accuracy scores of each round of the RFE.

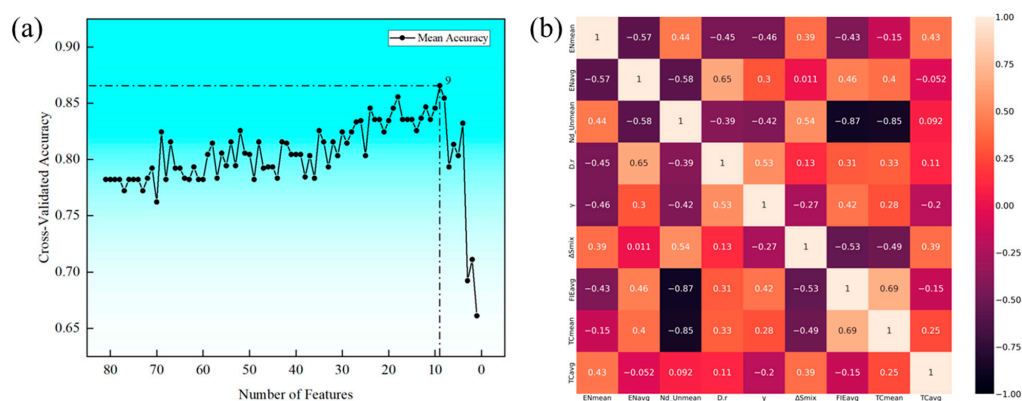


Figure 3. Results of RFE. (a) Features filtering by RFE. The dashed lines in the figure are the horizontal and vertical coordinates corresponding to the highest accuracy point; (b) PCC diagram of post-RFE feature.

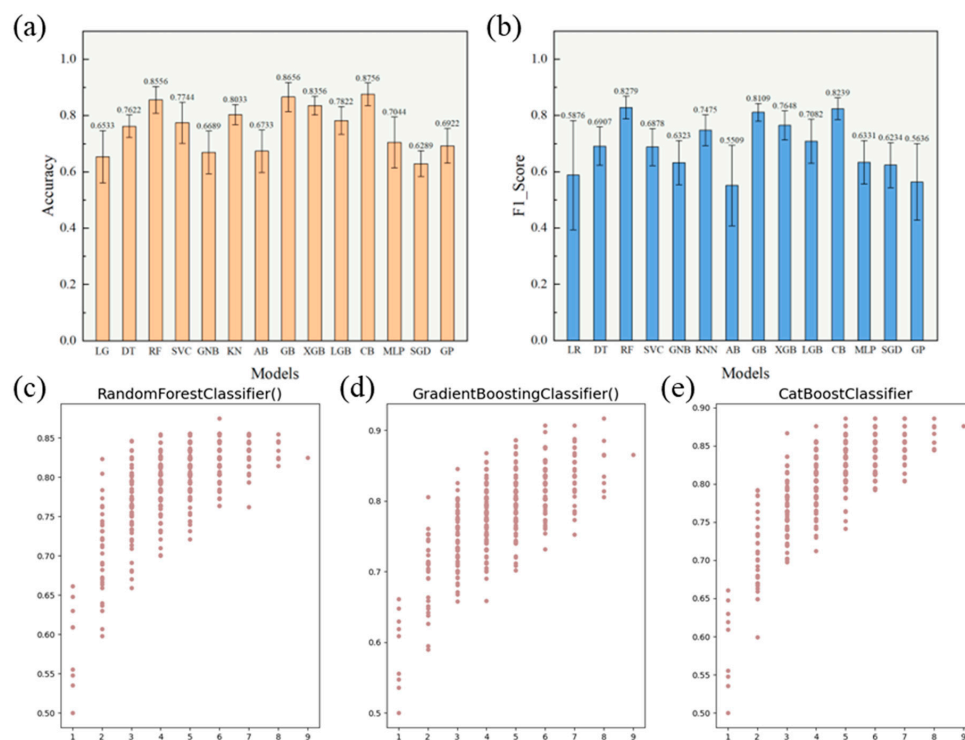
As can be seen from the figure, the accuracy of the model shows an increasing trend as the number of features decreases until only nine features remain, which means that eliminating features of lower importance can effectively improve the accuracy of the model. When features are further reduced, the accuracy of the model drops sharply, indicating that key features containing important information are deleted at this time. Therefore, it can be concluded that the unique subset of the nine features performs best. As a result, we obtained nine features related to atomic radius, valence electron concentration, mixing entropy, ionization energy, and thermal conductivity, as shown in Table 2.

Table 2. Feature candidates obtained from RFE.

Descriptor	Definition	Descriptor	Definition
EN _{mean}	Mean of electronegativity	ΔS_{mix}	Configuration entropy
EN _{avg}	Average deviation of electronegativity	FIE _{avg}	Average deviation of the first ionization energy
Nd_Un _{mean}	Mean of NdUnfilled	TC _{mean}	Mean of thermal conductivity
D.r	Radii local mismatch	TC _{avg}	Average deviation of thermal conductivity
γ	Radii gamma		

After RFE, we analyzed the Pearson correlation coefficients (PCC) between the remaining nine features. Figure 3b displays the PCC heat map, which illustrates the correlation of features through different shades of colors. In this figure, it can be observed that the PCC among the remaining nine features are all below 0.9, indicating that there is no strong correlation between these features.

Figure 4a,b shows the filtering results of the second step. We input the 9 features above into 14 different classification models, taking into account both accuracy and F1_Scores comprehensively. By comparing, three models (RF, GB, CB) are determined as model candidates for the next screening process in that they exhibit the highest accuracy and F1_Score.

**Figure 4.** Model selection and Best Subset Selection. (a) Accuracy of 14 models; (b) F1 score of 14 models; (c–e) Best Subset Selection via 3 models.

All possible combinations of the 10 features were given as the input of models and examined by three models. The comparison results are depicted in Figure 4c–e. After BSS, the accuracy of the three models has been significantly improved. However, the optimal feature subsets corresponding to each model were different, as shown in Table 3. To select the best model, classification reports were generated for the three models in

Figure 5, displaying precision, recall, and F1_Scores under different categories. To compare the models more intuitively, we created a histogram representing F1_Scores for different categories of each model, as shown in Figure 6.

Table 3. Feature candidates obtained from BSS.

Model	Best Subset of Features	Accuracy	F1_Score
Random Forest Classifier	EN _{mean} , EN _{vg} , Nd_Un _{mean} , γ , ΔS_{mix} , TC _{avg}	0.8633	0.7965
Gradient Boosting Classifier	EN _{mean} , EN _{avg} , D.r, γ , ΔS_{mix} , FIE _{avg} , TC _{mean} , TC _{avg}	0.9166	0.8923
Cat Boost Classifier	EN _{mean} , EN _{avg} , Nd_Un _{mean} , D.r, γ , ΔS_{mix} , FIE _{avg} , TC _{avg}	0.8855	0.8343

(a)	Precision	Recall	F1_Score	(b)	Precision	Recall	F1_Score
SS	0.72	0.59	0.65	SS	0.86	0.82	0.84
SS+IM	0.77	0.79	0.78	SS+IM	0.97	0.89	0.93
IM	0.90	0.97	0.94	IM	0.90	1.00	0.95
ave	0.80	0.78	0.79	ave	0.91	0.90	0.91

(c)	Precision	Recall	F1_Score
SS	0.84	0.73	0.78
SS+IM	0.87	0.87	0.87
IM	0.93	1.00	0.96
ave	0.88	0.87	0.87

Figure 5. The classification report of three models after adjustment. Here, the darker part means higher F1_Score. (a) Random Forest Classifier; (b) Gradient Boosting Classifier; (c) Cat Boost Classifier.

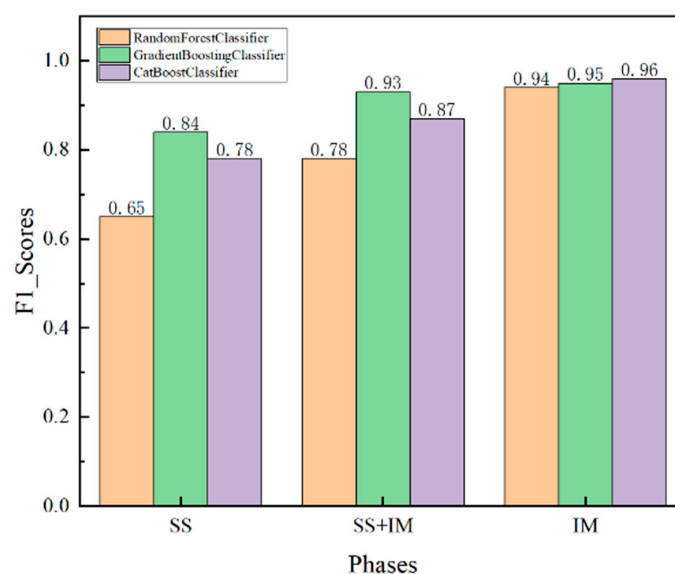


Figure 6. F1_Score for three models. (Here, orange, green, and purple are corresponding models for Random Forest Classifier, Gradient Boosting Classifier, and Cat Boost Classifier, respectively).

The three models have lower scores for SS prediction, possibly because there is less data about SS in the dataset. GB model's F1_Score is far greater than the other two models on the prediction of SS and SS and IM, which shows that the GB model has the best performance.

3.2. New Alloy Composition Design

In our work, it can only be reflected as yes or no in the prediction result because the classification problem is a qualitative problem. The prediction space contains a large number of SS components, so we need to take a method to guide us in choosing the appropriate alloy components for experimental verification. This study designed a resampling method, as shown in Figure 1b. The specific steps of resampling are as follows: we randomly selected 10% of the data from the dataset, combined it with the original dataset to create a new dataset, and then used this new dataset for predictions in the model. This process was repeated 1000 times.

We chose Al, Li, and Mg as the fixed elements for the five-element LHEAs. Additionally, we introduced two elements among Ti, Cu, and Zn, resulting in three alloy systems: AlLiMgZnCu, AlLiMgTiCu, and AlLiMgTiZn. The concentration range for each of the five elements is set to vary within 5 at% to 35 at%, with a step size of 1 at%. Finally, each alloy search space contained 553,401 LHEAs alloy data and the prediction of the phase in the search space was carried out by using the well-trained model. In these three systems, each system has a different probability of forming the SS phase. The AlLiMgZnCu system has the highest probability, reaching up to 86.4%, while the AlLiMgTiCu system has a maximum probability of 67.2%. But, the AlLiMgTiZn system's probability is below 36%. So, in the AlLiMgZnCu system, we selected the composition with the highest probability—Al₂₈Li₃₅Mg₁₅Zn₁₀Cu₁₂—for experimental validation.

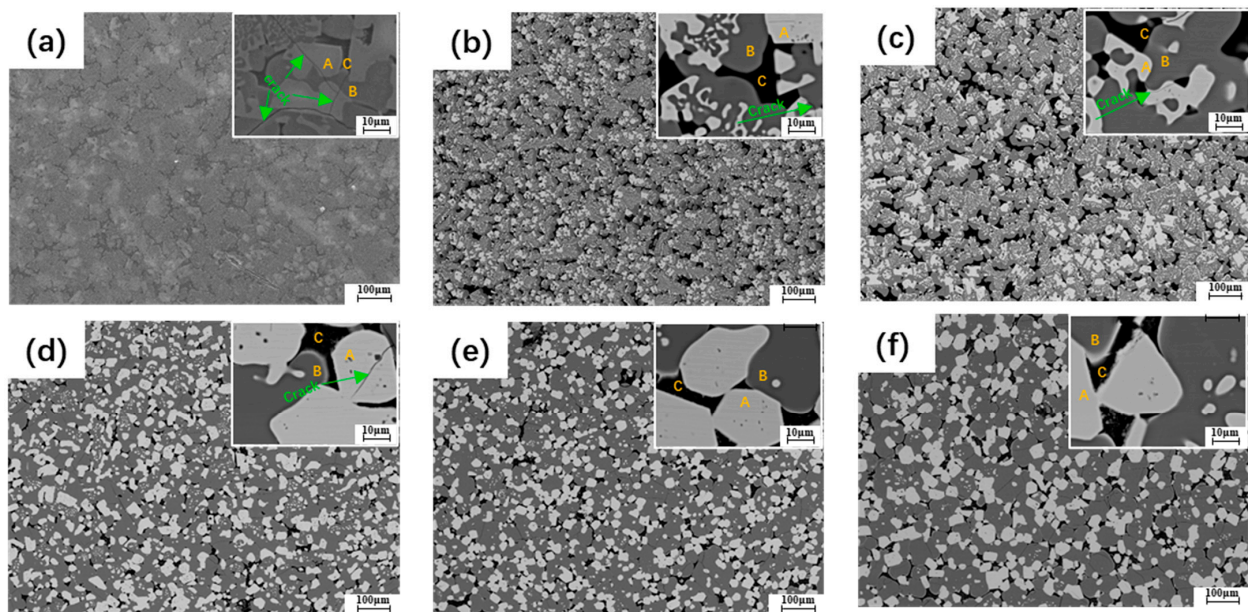
Why is the probability higher only in the AlLiMgZnCu system among these systems? The reason may be related to the mixing enthalpy between the elements. As mentioned in the criteria for the formation of HEAs phases, it is indicated that $-15 \text{ kJ/mol} < \Delta H_{mix} < 5 \text{ kJ/mol}$ is favorable for the formation of solid solution in HEAs. Guo et al. further narrowed down this criterion range to $-11.6 \text{ kJ/mol} < \Delta H_{mix} < 3.2 \text{ kJ/mol}$. These findings suggest that the absolute value of $\Delta H_{mix} = 0$ is favorable for solid solution formation. ΔH_{mix} is a parameter used to explain the interactions between constituent elements in multicomponent alloys. It can influence the solidification behavior of the alloy and the composition of phases after solidification. Such as a binary system at the equiatomic composition (A50B50) for the liquid phase, where the binary mixing enthalpy values ΔH_{i-j} (where ΔH_{i-j} is the mixing enthalpy of an A (*i*-th element)–B (*j*-th element)), the more negative ΔH_{i-j} indicates a stronger affinity between the elements. Conversely, more positive values suggest strong repulsion between the elements. The more negative the enthalpy of binary mixing between elements, the more favorable the formation of the amorphous phase. Whether the enthalpy of binary mixing enthalpy is more positive or negative, it is beneficial to the precipitation of intermetallic compounds [38,39]. On the contrary, the interactions between constituent elements are extremely weak in alloy systems where ΔH_{i-j} is zero or close to zero. In this case, liquid phase separation and the formation of amorphous phases are almost impossible, resulting in a greater tendency to form solid solutions [40]. The binary mixing enthalpies between aluminum (Al), lithium (Li), and magnesium (Mg) elements are close to zero. Hence, it is imperative to select elements that satisfy the following criteria: their mixing enthalpies with Al, Li, and Mg should be close to zero, and the mixing enthalpies between the chosen pairs of elements should also be close to zero. In this way, the alloy system composed of these elements is more likely to form the SS phase. In fact, only zinc (Zn) and copper (Cu) elements meet this criterion among common metals, as shown in Table 4. Hence, the high probability of forming the SS phase in the AlLiMgZnCu alloy system can be explained.

Table 4. Alloy design of Al-Li-Mg-X (X = Zn, Cu, Ti) alloy systems focusing on ΔH_{i-j} matrix adapted from [41].

	Al	Li	Mg	Zn	Cu	Ti
Al	-	−4	−2	1	−1	−30
Li	-	-	0	−7	−5	34
Mg	-	-	-	−4	−3	16
Zn	-	-	-	-	1	−15
Cu	-	-	-	-	-	−9
Ti	-	-	-	-	-	-

3.3. Experimental Verification

Figure 7 displays the microstructure of the $\text{Al}_{28}\text{Li}_{35}\text{Mg}_{15}\text{Zn}_{10}\text{Cu}_{12}$ LHAEs. From the SEM image in Figure 7a, it is observed that the alloy is mainly composed of continuous solid solution phases and discontinuous black second phases. According to phase statistics, the proportion of solid solutions achieved was 90.19%, which was in good agreement with the prediction result of machine learning. However, it is found that there are obvious contrast differences in the solid solution phase. In order to elucidate the reasons behind this, we conducted an EDS analysis, and Table 5 presents the elemental distribution in different regions of the alloy. It is evident that the brightness differences in regions A and B result from the distribution of elements, with region A having a higher concentration of Cu, while region B is dominated by Al and Zn elements. Region C is a Mg-rich second phase with minor amounts of Al and Zn.

**Figure 7.** SEM of $\text{Al}_{28}\text{Li}_{35}\text{Mg}_{15}\text{Zn}_{10}\text{Cu}_{12}$ alloys with different states. A, B, and C represent phases of different colors. (a) As-cast; (b) 400 °C/24 h; (c) 430 °C/24 h; (d) 460 °C/24 h; (e) 490 °C/24 h; (f) 520 °C/24 h.

In order to promote the redissolution of the second phase and eliminate elements' elongation, we have established different annealing heat treatment methods at different temperatures, including 400 °C/24 h, 430 °C/24 h, 460 °C/24 h, 490 °C/24 h, and 520 °C/24 h. The microstructures of the alloys under different annealing schedules are depicted in Figure 7b, which can be observed that the size and morphology of different phases have changed after annealing treatment. The proportion of the second phase decreases, but the contrast difference in the solid solution becomes more pronounced that phase separation occurs in the solid solution. When the annealing schedules are 400 °C/24 h and

430 °C/24 h, there is no significant change in the alloy's phase structure, which is attributed to the insufficient heat treatment temperature, leading to a lack of driving force for element diffusion. The Mg-rich phase in the alloy significantly decreases as the heat treatment temperature increases, which shows that the Mg-rich phase redissolved at 460 °C. Simultaneously, under Ostwald ripening [42], the same structure of different sizes in the solid solution phase merged. Further increasing the annealing temperature to 520 °C will not only lead to phase aggregation but also lead to phase growth which increases the average grain size. In addition, there are obvious cracks in the microstructure of the alloy, indicating that the LHEAs belong to brittle materials.

Table 5. Distribution of component elements (atomic fraction, %) in different regions of the $\text{Al}_{28}\text{Li}_{35}\text{Mg}_{15}\text{Zn}_{10}\text{Cu}_{12}$ light-weight high-entropy alloy.

Condition	Element Normal Component	Al	Mg	Zn	Cu
As-cast	A	28.2	32.8	8.4	30.6
	B	70.3	5.0	21.5	3.2
	C	5.4	89.2	4.5	0.9
400 °C/24 h	A	28.4	32.5	8.3	30.9
	B	69.1	7.2	21.0	2.7
	C	5.5	89.7	3.9	0.9
430 °C/24 h	A	28.7	32.4	8.2	30.7
	B	67.9	10.7	19.1	2.3
	C	6.0	89.3	4.0	0.7
460 °C/24 h	A	27.8	35.0	7.4	29.8
	B	69.6	6.5	20.3	3.6
	C	6.0	89.7	3.7	0.6
490 °C/24 h	A	26.2	39.3	6.1	28.4
	B	70.5	5.5	20.5	3.5
	C	5.6	89.7	4.0	0.7
520 °C/24 h	A	26.4	39.2	6.3	28.1
	B	70.2	5.7	20.8	3.3
	C	5.3	89.4	4.6	0.7

Although the solid solution proportion of the predicted alloy is as high as 90%, this also fulfilled the goal of machine learning. However, the solid solution of this alloy is obviously a two-phase solid solution, and we essentially expect to obtain an alloy dominated by a single-phase solid solution. So, we continued with the following exploration.

For the two solid solution structures with different compositions, we, respectively, designed $\text{Al}_{24}\text{Li}_{15}\text{Mg}_{26}\text{Zn}_9\text{Cu}_{26}$ (1#), $\text{Al}_{21}\text{Li}_{25}\text{Mg}_{24}\text{Zn}_7\text{Cu}_{23}$ (2#), and $\text{Al}_{61}\text{Li}_{15}\text{Mg}_5\text{Zn}_{16}\text{Cu}_3$ (3#) LHEAs, whose purpose is to improve the ability of the alloy to form a single solid solution by adjusting the alloy composition. The SEM and EDS results of the alloys are shown in Figure 8, and the elemental contents of different phases are listed in Table 6. Clearly, the 1# alloy achieved a phase structure predominantly composed of solid solution phases, with the proportion of solid solution phases exceeding 80%. However, there still existed a small amount of intermetallic compounds. The difference between the 1# and 2# alloys differs in the Li, which makes their structures distinguished. It can be found that the more Li content, the easier it is to form intermetallic compounds. Compared with $\text{Al}_{28}\text{Li}_{35}\text{Mg}_{15}\text{Zn}_{10}\text{Cu}_{12}$ alloys, the elemental content in the solid solution phase of 1# and 2# alloys has changed greatly. The most obvious elements are Mg and Cu, the content of Mg decreases clearly, and the atomic ratio of Cu increases from 30 at% to 62 at%.

In conclusion, increasing the Cu content and reducing the Mg and Li can be beneficial to improve the tendency to form the solid solution phase. This phenomenon may be attributed to the large atomic radius of Mg and Li elements, which is not conducive to the formation of solid solution phases, and the lack of d-orbital electrons in the electron shell structure hinders higher-order hybridization with other elements.

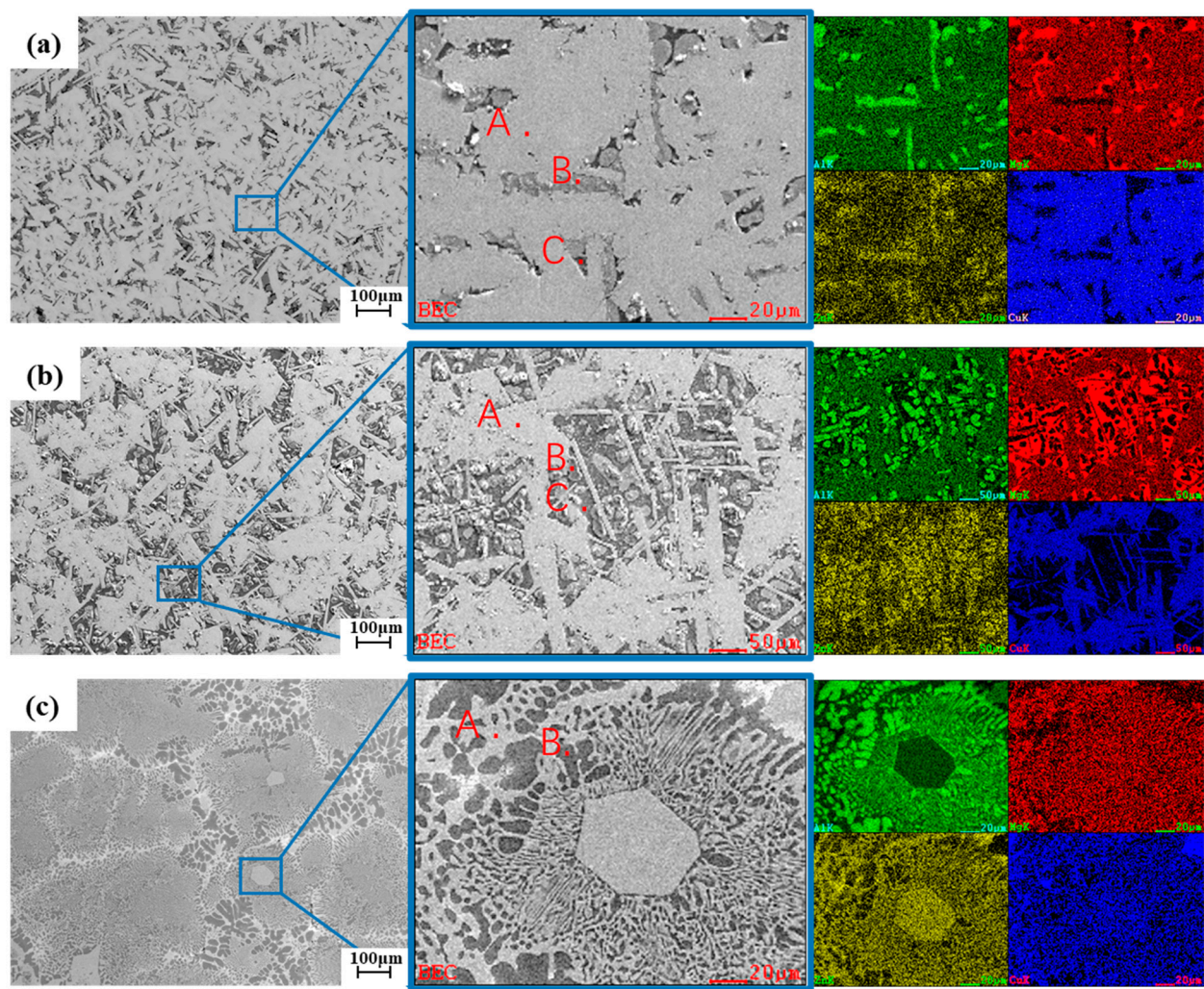


Figure 8. SEM images and EDS maps of light-weight high-entropy alloy. In the picture, region A is the white phase, region B is the gray phase, and region C is the black phase rich in magnesium. (a) $\text{Al}_{24}\text{Li}_{15}\text{Mg}_{26}\text{Zn}_9\text{Cu}_{26}$; (b) $\text{Al}_{21}\text{Li}_{25}\text{Mg}_{24}\text{Zn}_7\text{Cu}_{23}$; (c) $\text{Al}_{61}\text{Li}_{15}\text{Mg}_5\text{Zn}_{16}\text{Cu}_3$.

Table 6. Distribution of component elements (atomic fraction, %) in different regions of the $\text{Al}_{24}\text{Li}_{15}\text{Mg}_{26}\text{Zn}_9\text{Cu}_{26}$, $\text{Al}_{21}\text{Li}_{25}\text{Mg}_{24}\text{Zn}_7\text{Cu}_{23}$, $\text{Al}_{61}\text{Li}_{15}\text{Mg}_5\text{Zn}_{16}\text{Cu}_3$ light-weight high-entropy alloy.

Number	Alloy	Element Normal Component	Al	Mg	Zn	Cu
1#	$\text{Al}_{24}\text{Li}_{15}\text{Mg}_{26}\text{Zn}_9\text{Cu}_{26}$	A	14.3	15.3	8.4	62.0
		B	52.2	5.2	35.0	7.6
		C	1.3	85.3	10.4	3.0
2#	$\text{Al}_{21}\text{Li}_{25}\text{Mg}_{24}\text{Zn}_7\text{Cu}_{23}$	A	13.2	15.0	8.8	63.0
		B	51.4	2.3	34.9	11.4
		C	2.6	83.5	12.3	1.6
3#	$\text{Al}_{61}\text{Li}_{15}\text{Mg}_5\text{Zn}_{16}\text{Cu}_3$	A	40.9	6.4	43.3	9.4
		B	86.0	0.9	11.6	1.5

Moreover, the 3# alloy has a network interwoven structure of white phase and gray phase, and the microstructure is coarse at the grain boundaries. According to EDS analysis, the white phase has a higher content of Mg, Zn, and Cu elements, while the gray phase is the Al-rich phase. It is worth mentioning that the network structure of the 3# alloy may contribute to the improvement of its properties.

Similarly, we selected 1# and 3# alloys with relatively simple microstructures for annealing heat treatment to improve their microstructure. Referring to the aforementioned annealing treatment, we selected temperatures of 430 °C/24 h, 460 °C/24 h, and 490 °C/24 h, and the alloy microstructures are shown in Figure 9. For the 1# alloy, there was no significant change at the annealing temperature of 430 °C/24 h. As the temperature increased to 460 °C, the alloy structure began to resolve. When the temperature is 490 °C, all of the Mg-rich phase and some of the gray phase redissolved, resulting in a solid solution phase proportion exceeding 90%. Furthermore, the structure of the 3# alloy only experienced growth after heat treatment.

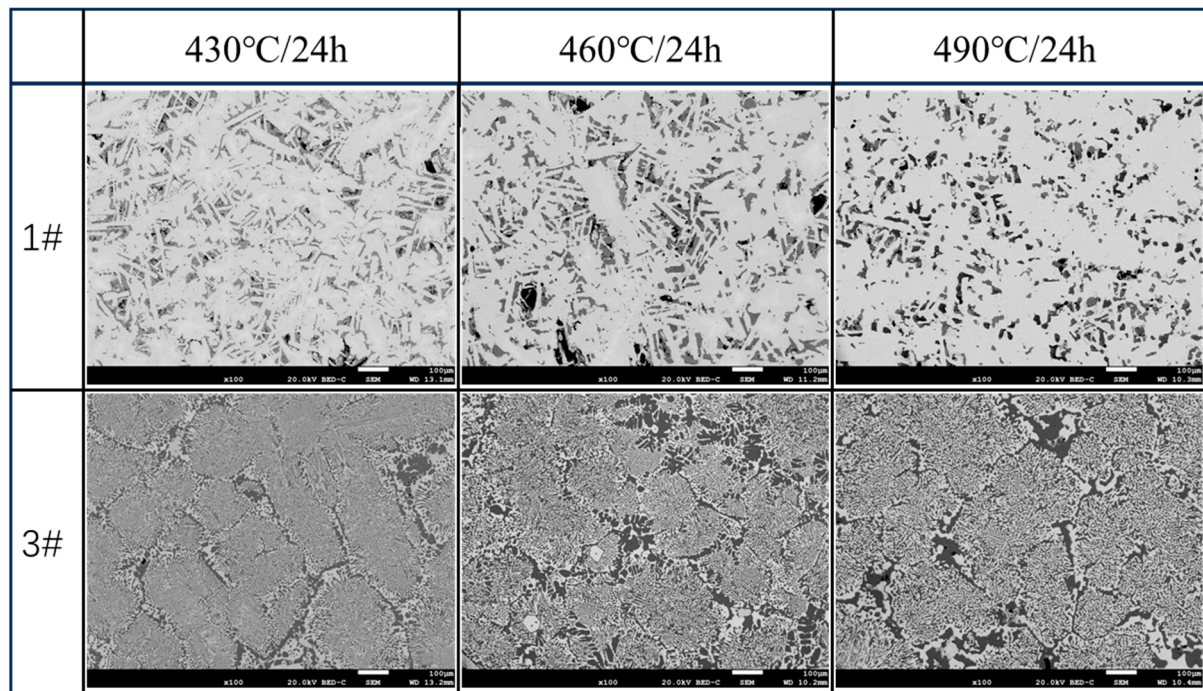


Figure 9. The microstructure of the designed alloy in as-cast and different heat treatment states.

4. Conclusions

This study aims to establish a machine learning model to discover LHEAs with simple solid solution phases. (1) Through a series of feature engineering and model selection steps, the well-trained Gradient Boosting Classifier model achieved a high accuracy of 0.9166 and an F1_Score of 0.8923. The features with high correlation with the phases of alloys include EN_{mean} , EN_{avg} , $D.r$, γ , ΔS_{mix} , FIE_{avg} , TC_{mean} , and TC_{avg} , which is related to electronic structure and electron density. (2) Using machine learning, we designed $Al_{28}Li_{35}Mg_{15}Zn_{10}Cu_{12}$ light-weight high-entropy alloys, with over 90% solid solution phase, demonstrating the effectiveness of the machine learning predictions. However, the solid solution of the alloy is a two-phase solid solution. (3) Combining machine learning prediction results and expert advice, we adjusted the composition to design $Al_{24}Li_{15}Mg_{26}Zn_9Cu_{26}$ alloy, which has a single solid solution, and the content of a single solid solution after heat treatment is also as high as 90%.

Overall, although the predicted LHEAs still contain intermetallic compounds, the alloys are mainly dominated by continuous a single solid solution, which provides a direction for future exploration of LHEAs. It is worth mentioning that this research reinforces the idea that expert experience plays a significant role in discovering materials.

Author Contributions: Conceptualization, H.Y., Z.L., Y.Z. and B.X.; methodology, Z.L., Y.L., X.L. and Y.Z.; validation, S.L.; formal analysis, S.L. and Y.L.; resources, Y.L. and X.L.; data curation, S.L. and R.L.; writing—review and editing, R.L., H.Y., Y.Z. and B.X.; visualization, S.L.; supervision, H.Y., Y.Z.

and B.X.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the Science and Technology Innovation Fund Project of GRIMAT Engineering Institute Co., Ltd., Beijing, China.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: All authors were employed by the companies GRINM Group Co., Ltd. and General Research Institute for Nonferrous Metals. Authors Shenglong Li, Rui Liu, Hongwei Yan, Yanan Li, Xiwu Li and Yongan Zhang were employed by the company GRIMAT Engineering Institute Co., Ltd. All the authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Xue, H.; Yang, C.; De, G.F.; Zhang, P.; Zhang, J.Y.; Chen, B.; Liu, F.Z.; Peng, Y.; Bian, J.J.; Liu, G.; et al. Highly stable coherent nanoprecipitates via diffusion-dominated solute uptake and interstitial ordering. *Nat. Mater.* **2023**, *22*, 434–441. [\[CrossRef\]](#)
2. Cantor, B. Multicomponent and high entropy alloys. *Entropy* **2014**, *16*, 4749–4768. [\[CrossRef\]](#)
3. Cantor, B.; Chang, I.T.H.; Knight, P.; Vincent, A.J.B. Microstructural development in equiatomic multicomponent alloys. *Mater. Sci. Eng. A* **2004**, *375*, 213–218. [\[CrossRef\]](#)
4. Yeh, J.W.; Chen, S.K.; Lin, S.J.; Gan, J.Y.; Chin, T.S.; Shun, T.T.; Tsau, C.H.; Chang, S.Y. Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes. *Adv. Eng. Mater.* **2004**, *6*, 299–303. [\[CrossRef\]](#)
5. Ozalp, A.; Okuyucu, C.; Koc, B.; El, A.O.; Aydogan, E. Development and directed energy deposition of high strength $\text{Hf}_5\text{Mo}_{15}\text{Nb}_{35}\text{Ta}_{25}\text{Ti}_{20}$ refractory high entropy alloys. *Mater. Charact.* **2024**, *209*, 113679. [\[CrossRef\]](#)
6. Geng, Z.W.; Chen, C.; Song, M.; Luo, J.R.; Chen, J.X.; Li, R.D.; Zhou, K.C. High strength $\text{Al}_{0.7}\text{CoCrFeNi}_{2.4}$ hypereutectic high entropy alloy fabricated by laser powder bed fusion via triple-nanoprecipitation. *J. Mater. Sci. Technol.* **2024**, *187*, 141–155. [\[CrossRef\]](#)
7. Senkov, O.N.; Senkova, S.V.; Woodward, C. Effect of aluminum on the microstructure and properties of two refractory high-entropy alloys. *Acta Mater.* **2014**, *68*, 214–228. [\[CrossRef\]](#)
8. Senkov, O.N.; Senkova, S.V.; Miracle, D.B.; Woodward, C. Mechanical properties of low-density, refractory multi-principal element alloys of the Cr–Nb–Ti–V–Zr system. *Mater. Sci. Eng. A* **2013**, *565*, 51–62. [\[CrossRef\]](#)
9. Kalali, D.G.; Antharam, S.; Hasan, M.; Karthik, P.S.; Phani, P.S.; Rao, K.B.S.; Rajulapati, K.V. On the origins of ultra-high hardness and strain gradient plasticity in multi-phase nanocrystalline MoNbTaTiW based refractory high-entropy alloy. *Mater. Sci. Eng. A* **2021**, *812*, 141098. [\[CrossRef\]](#)
10. Luan, H.W.; Shao, Y.; Zhang, X.R.; Zhu, J.; Yang, S.H.; Si, J.J.; Li, J.F.; Yao, K.F. Effect of TiC Addition on the High-Temperature Hardness and the Carbide Stability in $\text{Al}_{0.2}\text{CoCrFeNi}_{1.5}\text{Ti}$ High-Entropy Alloy. *Russ. J. Non-Ferr. Met.* **2022**, *62*, 708–715. [\[CrossRef\]](#)
11. Olga, S.; Nataliya, S.; Ahmad, M.O.; Evgeny, T. $\text{Al}_{0.25}\text{CoCrFeNiSi}_{0.6}$ high entropy alloy with high hardness and improved wear resistance. *Mater. Lett.* **2022**, *328*, 133190.
12. Duan, J.C.; Wang, M.L.; Huang, R.; Miao, J.W.; Lu, Y.P.; Wang, T.M.; Li, T.J. A novel high-entropy alloy with an exceptional combination of soft magnetic properties and corrosion resistance. *Sci. China Mater.* **2022**, *66*, 772–779. [\[CrossRef\]](#)
13. Gu, X.H.; Huang, Q.Y.; Chen, J.B.; Hu, H.S.; Sun, Q.Q.; Wu, L.K.; Cao, F.H. Optimizing corrosion resistance of equiatomic AlCoCrFeNi high entropy alloys via heat treatment. *J. Alloys Compd.* **2023**, *968*, 172091. [\[CrossRef\]](#)
14. Lin, C.; Meng, L.; Tan, M.T.; Ke, L.S.; Qi, T. Microstructures and properties of $\text{Cr}_x\text{-FeNiCu}_{0.5}\text{Ti}_{0.5}$ high-entropy alloys for corrosion resistance. *Intermetallics* **2023**, *153*, 107781. [\[CrossRef\]](#)
15. Zhang, W.; Jiang, W.; Wu, S.S.; Tao, S.T.; Zhu, B.H.; Guo, S.L. Microstructural and Mechanical Evaluation of NbMoTiVSi_x ($x = 0, 0.25$) Refractory High-Entropy Alloys at High Temperature. *Adv. Eng. Mater.* **2022**, *24*, 2101248. [\[CrossRef\]](#)
16. Sun, J.; Zhao, W.X.; Yan, P.; Li, S.Y.; Dai, Z.C.; Jiao, L.; Qiu, T.Y.; Wang, X.B. High temperature tensile properties of as-cast and forged CrMnFeCoNi high entropy alloy. *Mater. Sci. Eng. A* **2022**, *850*, 143570. [\[CrossRef\]](#)
17. Zhou, Y.H.; Zhang, J.Y.; Zhang, J.; Yao, X.Y.; Luan, J.H.; Li, Q.; Liu, S.F.; Xiao, B.; Ju, J.; Zhao, S.J.; et al. A strong-yet-ductile high-entropy alloy in a broad temperature range from cryogenic to elevated temperatures. *Acta Mater.* **2024**, *268*, 119770. [\[CrossRef\]](#)
18. King, D.J.M.; Middleburgh, S.C.; McGregor, A.G.; Cortie, M.B. Predicting the formation and stability of single phase high-entropy alloys. *Acta Mater.* **2016**, *104*, 172–179. [\[CrossRef\]](#)
19. Jia, Y.F.; Jia, Y.D.; Wu, S.W.; Ma, X.D.; Wang, G. Novel ultralight-weight complex concentrated alloys with high strength. *Materials* **2019**, *12*, 1136. [\[CrossRef\]](#)
20. Yang, X.; Chen, S.Y.; Cotton, J.D.; Zhang, Y. Phase stability of low-density, multiprincipal component alloys containing aluminum, magnesium, and lithium. *JOM* **2014**, *66*, 2009–2020. [\[CrossRef\]](#)

21. Li, R.; Gao, J.C.; Fan, K. Study to microstructure and mechanical properties of Mg containing high entropy alloys. *Mater. Sci. Forum* **2010**, *650*, 265–271. [\[CrossRef\]](#)
22. Li, Y.; Guo, W.L. Machine-learning model for predicting phase formations of high-entropy alloys. *Phys. Rev. Mater.* **2019**, *3*, 095005. [\[CrossRef\]](#)
23. Nong, Z.S.; Zhu, J.C.; Cao, Y.; Yang, X.W.; Lai, Z.H.; Liu, Y. Stability and structure prediction of cubic phase in as cast high entropy alloys. *Mater. Sci. Technol.* **2014**, *30*, 363–369. [\[CrossRef\]](#)
24. Huang, W.; Martin, P.; Zhuang, H.L. Machine-learning phase prediction of high-entropy alloys. *Acta Mater.* **2019**, *169*, 225–236. [\[CrossRef\]](#)
25. Ramalhete, P.S.; Senos, A.M.R.; Aguiar, C. Digital tools for material selection in product design. *Mater. Des.* **2009**, *31*, 2275–2287. [\[CrossRef\]](#)
26. Feng, R.; Gao, M.C.; Lee, C.; Mathes, M.; Zuo, T.T.; Chen, S.Y.; Hawk, J.A.; Zhang, Y.; Liaw, P.K. Design of light-weight high-entropy alloys. *Entropy* **2016**, *18*, 333. [\[CrossRef\]](#)
27. Sanchez, J.M.; Vicario, I.; Albizuri, J.; Guraya, T.; Koval, N.E.; Garcia, J.C. Compound formation and microstructure of as-cast high entropy aluminums. *Metals* **2018**, *8*, 167. [\[CrossRef\]](#)
28. Kumar, A.; Gupta, M. An insight into evolution of light weight high entropy alloys: A review. *Metals* **2016**, *6*, 199. [\[CrossRef\]](#)
29. Zhang, Y.; Zhou, Y.J.; Lin, J.P.; Chen, G.L.; Liaw, P.K. Solid-solution phase formation rules for multi-component alloys. *Adv. Eng. Mater.* **2008**, *10*, 534–538. [\[CrossRef\]](#)
30. Xiong, J.; Shi, S.Q.; Zhang, T.Y. A machine-learning approach to predicting and understanding the properties of amorphous metallic alloys. *Mater. Des.* **2020**, *187*, 108378. [\[CrossRef\]](#)
31. Wen, C.; Zhang, Y.; Wang, C.X.; Xue, D.Z.; Bai, Y.; Antonov, S.; Dai, L.H.; Lookman, T.; Su, Y.J. Machine learning assisted design of high entropy alloys with desired property. *Acta Mater.* **2019**, *170*, 109–117. [\[CrossRef\]](#)
32. Friedman, L.; Komogortsev, O.V. Assessment of the effectiveness of seven biometric feature normalization techniques. *IEEE. Trans. Inf. Forensics Secur.* **2019**, *14*, 2528–2536. [\[CrossRef\]](#)
33. Cai, J.; Luo, J.W.; Wang, S.L.; Yang, S. Feature selection in machine learning: A new perspective. *Neurocomputing* **2018**, *300*, 70–79. [\[CrossRef\]](#)
34. Eswara, K.R.; Appa, G.R. Retraction Note: Ensemble learning with recursive feature elimination integrated software effort estimation: A novel approach. *Evol. Intell.* **2023**, *16*, 719.
35. Li, H.Y.; Li, X.W.; Li, Y.N.; Xiao, W.; Wen, K.; Li, Z.H.; Zhang, Y.A.; Xiong, B.Q. Machine learning assisted design of aluminum-lithium alloy with high specific modulus and specific strength. *Mater. Des.* **2023**, *225*, 111483. [\[CrossRef\]](#)
36. Kong, W.K.X.; Zhu, J.; Bi, S.Z.; Huang, L.T.; Wu, P.; Zhu, S.J. Adaptive best subset selection algorithm and genetic algorithm aided ensemble learning method identified a robust severity score of COVID-19 patients. *iMeta* **2023**, *2*, 126. [\[CrossRef\]](#)
37. Matin, S.M. Improving the prediction of wind speed and power production of SCADA system with ensemble method and 10-fold cross-validation. *Case Stud. Chem. Environ. Eng.* **2023**, *8*, 100351.
38. Inoue, A. Stabilization of metallic supercooled liquid and bulk amorphous alloys. *Acta Mater.* **2000**, *48*, 279–306. [\[CrossRef\]](#)
39. Kündig, A.A.; Ohnuma, M.; Ping, D.H.; Ohkubo, T.; Hono, K. In situ formed two-phase metallic glass with surface fractal microstructure. *Acta Mater.* **2004**, *52*, 2441–2448. [\[CrossRef\]](#)
40. Nagase, T.; Terayama, A.; Nagaoka, T.; Fuyama, N.; Sakamoto, T. Alloy Design and Fabrication of Ingots of Al–Mg–Li–Ca Light-Weight Medium Entropy Alloys. *Mater. Trans.* **2020**, *61*, 1369–1380. [\[CrossRef\]](#)
41. Takeuchi, A.; Inoue, A. Classification of Bulk Metallic Glasses by Atomic Size Difference, Heat of Mixing and Period of Constituent Elements and Its Application to Characterization of the Main Alloying Element. *Mater. Trans.* **2005**, *46*, 2817–2829. [\[CrossRef\]](#)
42. Kang, M.K.; Kim, D.Y.; Hwang, N.M. Ostwald ripening kinetics of angular grains dispersed in a liquid phase by two-dimensional nucleation and abnormal grain growth. *J. Eur. Ceram. Soc.* **2002**, *22*, 603–612. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.