Article

Not peer-reviewed version

Machine learning-driven Design of High-elastocaloric NiTi-based Shape Memory Alloys

Yingyu Gao , Yunfeng Hu , Xinpeng Zhao , Yang Liu , Haiyou Huang * , Yanjing Su

Posted Date: 21 August 2024

doi: 10.20944/preprints202408.1415.v1

Keywords: machine learning; shape memory alloys; elastocaloric effect; phase transformation-induced entropy change; processing parameters

Article

Machine Learning-Driven Design of High-Elastocaloric NiTi-Based Shape Memory Alloys

Yingyu Gao [1,2], Yunfeng Hu [1], Xinpeng Zhao [1], Yang Liu [1,3], Haiyou Huang [1,3,4,*] and Yanjing Su [1,3,4]

[1] Institute for Advanced Materials and Technology, University of Science and Technology Beijing, Beijing 100083, China
[2] Xiamen Hithium Energy Storage Technology Co., Ltd.
[3] Beijing Key Laboratory of Materials Genome Engineering, University of Science and Technology Beijing, Beijing 100083, China
[4] Beijing Advanced Innovation Center for Materials Genome Engineering, University of Science and Technology Beijing, Beijing 100083, China
* Correspondence: huanghy@mater.ustb.edu.cn; Tel.: +86-010-6233-3884

Abstract: In recent years, due to the detrimental impact of traditional gas-liquid refrigerants on the environment, the pursuit of solid-state refrigeration technology based on elastocaloric effect, as a promising sustainable alternative, has been promoted. However, the traditional trial-and-error method is inefficient and difficult to meet the material application requirements. In this work, machine learning (ML) was employed to accelerate the development of NiTi-based shape memory alloy (SMA) with excellent elastocaloric effect. By means of active learning, we identified 9 novel NiTi-based SMAs in four iterations, the phase transformation-induced entropy change (ΔS) of which are as high as over 90 J/kg·K-1. Moreover, the model exhibits good interpretability, allowing us to understand the relevance between the target and features from the aspect of mechanism.

Keywords: machine learning; shape memory alloys; elastocaloric effect; phase transformation-induced entropy change; processing parameters

1. Introduction

Traditional refrigeration relies on gas-liquid compression technology, utilizing common refrigerants like chlorofluorocarbon (CFC) and hydrochlorofluorocarbon (HCFC), which pose a serious threat to the ozone layer. While some refrigerants such as hydrofluorocarbon (HFC) don't harm the ozone layer, they contribute significantly to the greenhouse effect[1]. Moreover, conventional refrigeration systems are plagued by issues such as excessive noise levels and high energy consumption. Refrigeration alone consumes over 15% of global electricity, a figure anticipated to escalated with economic growth[2]. In response, researchers are actively exploring alternative cooling materials and technologies.

Various solid-state refrigeration technologies based on different mechanisms, such as the magnetocaloric effect[3], electrocaloric effect[4], elastocaloric effect[5], have been proposed as alternatives due to their environmentally friendly. Particularly, the elastocaloric effect, specifically based on shape memory alloys (SMAs), has shown its unique promise[6], on account of a coefficient of performance (COP) exceeding 80%[7]. Notably, NiTi-based SMAs outperform Cu-based, Fe-based and ferromagnetic SMAs in terms of latent heat of phase change and elastomeric effects. As a result, prototype chillers utilizing NiTi-based SMAs have been successfully reported[8,9].

The composition[11–18] and preparative process[19,20] of NiTi-based SMAs wield a significant influence on phase transition temperature, precipitation, and microstructure. These factors, consequently, play a crucial role in affecting the ΔS of elastocaloric effect. In the case of multielement NiTi alloys, the variables concerning the type and content of the alloying elements span a wide range.

 doi:10.20944/preprints202408.1415.v1

Moreover, common heat treatment processes for NiTi-based SMAs encompass homogenization, solution, and aging treatments[21]. Each step in these processes have a broad range of selectable heating temperature, time and cooling method. Given the extensive search space involved in combined composition and process parameters, machine learning (ML) has proven to be a highly effective tool in materials research and engineering practice[22–31].

In this study, we developed machine learning models to predict the entropy change of phase transition (ΔS) in NiTi-based SMAs using both physical characteristics and heat treatment process parameters. Adopting an active learning strategy[32,33], we iteratively optimized new materials that is 9 new NiTi-based SMA with ΔS greater than 90 J/Kg·K-1 after four iterations. We further optimized the corresponding heat treatment process parameters for these new alloys and conducted experimental verification. Finally, we delved into the physical interpretability of our ML models.

2. Methods

2.1. Datasets and Features

The first ML model for prediction of ΔS constructed using physical features (denoted as M1) contain 137 data points taken from published literatures, and including the ΔS values and 60 physicochemical characteristics. The second ML model (denoted as M2) constructed using physical features combined with heat treatment process parameters has 200 data points, including the ΔS values, 20 physical features and 9 heat treatment process parameters, as shown in Table S1 in Supplementary Materials. These physicochemical characteristics encompass information at the electron, atomic and lattice levels. For NiTi-based SMAs, generally, there are three heat treatment stages: homogenization treatment (I), solution treatment (II) and aging treatment (III). Common cooling methods include water quenching (WQ), oil quenching (OQ), air cooling (AC), and cooling in the furnace (FC). To enable the ML models to handle string-type data[34], we used label encoding to encode WQ, OQ, AC and FC as 4, 3, 2, and 1 respectively. Some of the data have only one or two of the three ways of heat treatment for the SMAs. To make more efficient use of the limited dataset, we complemented the missing solution and aging heat treatment parameters as heating at room temperature of 293 K for 0 min with air cooling, and the homogenization was also added, as shown in Table 1.

Table 1. Filling data for missing value of heat treatment process parameters.

Heat treatment process parameter	Filling value
HT_{I}（K）	1123 / 1223 / 1273 / 1323
Ht_{I}（mins）	4320
W_{I}	WQ / OQ / AC / FC
HT_{II}（K）	293
Ht_{II}（mins）	0
W_{II}	AC
HT_{III}（K）	293
Ht_{III}（mins）	0
W_{III}	AC

2.2. Experiments

In this work, our experiments focused on the exploration and extrapolation of the composition and process of NiTi-based binary and ternary SMAs, which guided by machine learning models and are used for model validation and iterative learning. The ingots of SMAs were prepared using 99.99% pure nickel, titanium and other alloying metals by arc melting under argon atmosphere. The specimens corresponding to the predicted results from M1 weren't heat treated, and the specimens corresponding to the predicted results of M2 were heat-treated according to the predicted process parameters. Samples with a size of 5 *mm*×5 *mm*×3 *mm* and Φ3 mm×1.5 mm cut from the ingots were

 doi:10.20944/preprints202408.1415.v1

ground and polished. And then their elemental contents were verified using an energy dispersive spectrometer (EDS) and phase transformation temperatures and latent heat ΔH were measured using a differential scanning calorimetry (DSC).

3. Results and Discussion

3.1. ML Model Selection and Evaluation

The Figures 1(a) and 2(a) illustrate the Pearson correction between features for M1 and M2, respectively. After removing the strongly correlated features with a Pearson correlation coefficients threshold of 0.8 for M1, we retained 13 features (Table S2) for details). Similarly, 10 physicochemical features and 9 heat treatment process features were preserved for M2 (pink-marked in Table S1).

Notably, the temperature, time and quenching methods of three heat treatment processes were independent and only related to the experimenter's setup. Consequently, these factors were not considered in the feature screening process.

Subsequently, we evaluated the R^2 and RMSE of various regression models using the standardized and pre-screened dataset. The optimal model, selected based on the largest R^2 and smallest RMSE, from Support Vector Machine with radial kernel (SVR. rbf), Random Forest (RF), Gradient Boosting (GBR), Ridge Regression (Ridge), Least absolute shrinkage and selection operator (LASSO), Elastic Network (EN), and Linear regression model (Lin). We finetuned the hyperparameters of these models using leave-out cross-validation[35] and Grid Search or Random Search.

As shown in Figure 1(b) and 1(c), for M1, the top performer was SVR. rbf. After a more exhaustive features screening, we identified 4 key physical features for constructing M1: $\Delta S = f_1(aven, en, Nsvalence, Modulus, compression)$. The resulting R^2 for M1 was 0.860, and the RMSE was 8.472, which were obtained according to Eq. (1) and (2).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{\sum_{i=1}^{n}(y_i - \bar{y})} \quad (1)$$

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

where $\hat{y}_i$ is the predicted value of the output of each cluster, y_i is the experimental data, and $\bar{y}$ is the mean value of y.

Similarly, for M2, the best-performing model was RF, with $\Delta S = f_2$(aw, dor, rcov, Valence, $HT_{I,II,III}$, $Ht_{I,II,III}$, $W_{I,II,III}$). The R2 of M2 was 0.746 and the RMSE was 10.422. However, the generally accepted threshold for better interpretability in regression model is R2 over 0.8. Therefore, to enhance the R2 of M2, we explored ensemble learning[36–38] (Boosting, Bagging and Stacking), cluster learning (K-Means[39] and DBSCAN[40]) and Neural Networks[41]. The segmentation fitting method showed superior results, and donated as M3. Using M1 as a reference, we clustered points with similar heat treatment process parameters into four clusters with 45,50,49,56 data points, respectively. The R2 and RMSE of the optimal models for these clusters were showcased in Figure 3 and Table 2. The final R2 for M3 was 0.866, with an RMSE of 7.544. The divide-and-conquer approach demonstrated better fitness and model performance compared to the traditional RF model M2.

 doi:10.20944/preprints202408.1415.v1

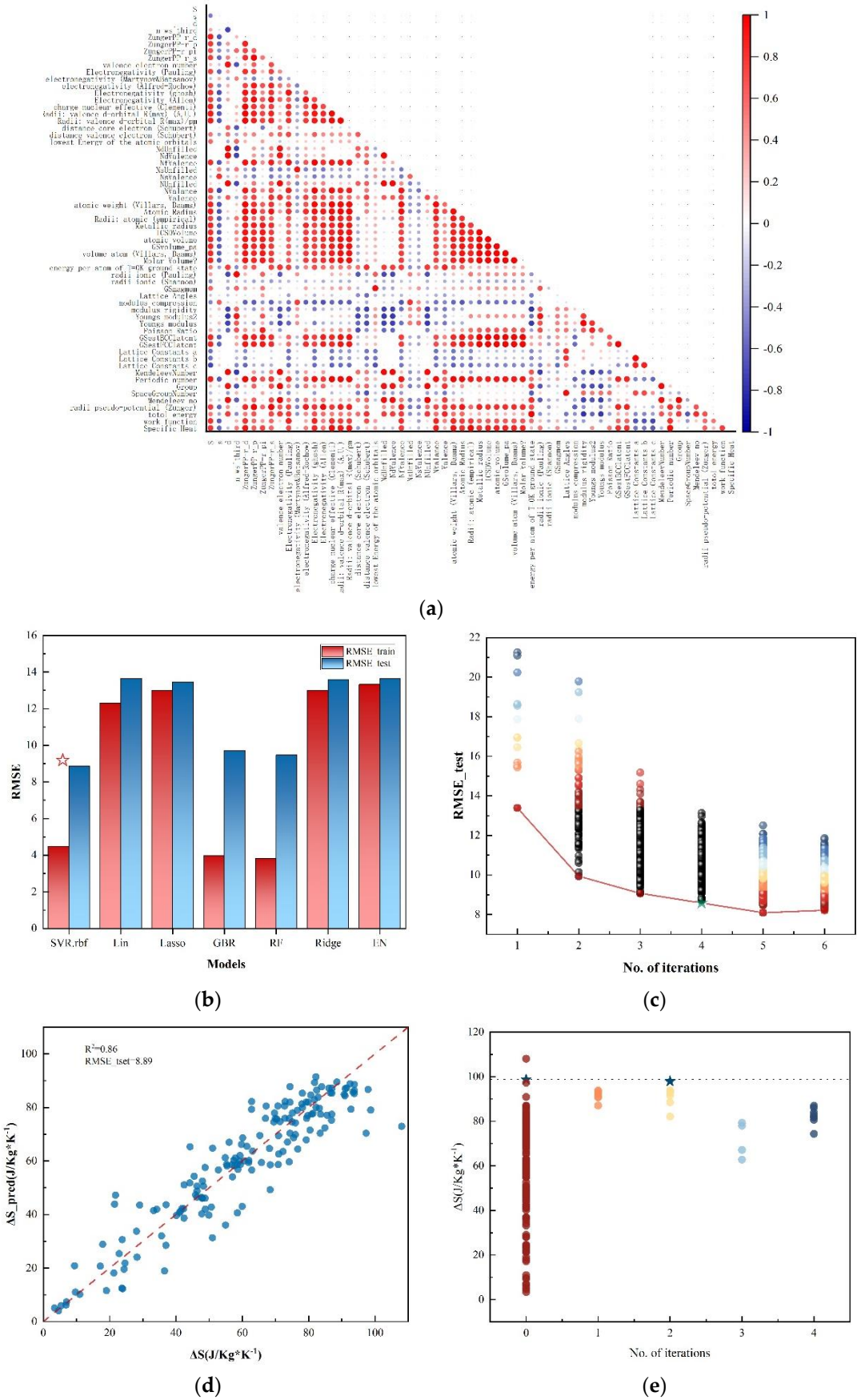

Figure 1. Characterization and model selection for the physical and chemical characteristic model (M1). (a) Pearson correlation matrix for the initial 60 descriptors. (b) Model selection based on leave-out cross-validation. (c) Cross-validation error from SVR.rbf model containing a subset of the retained

 doi:10.20944/preprints202408.1415.v1

descriptors. (d) The prediction performance of M1. (e) Results of 4 iterative of active learning based on M1.

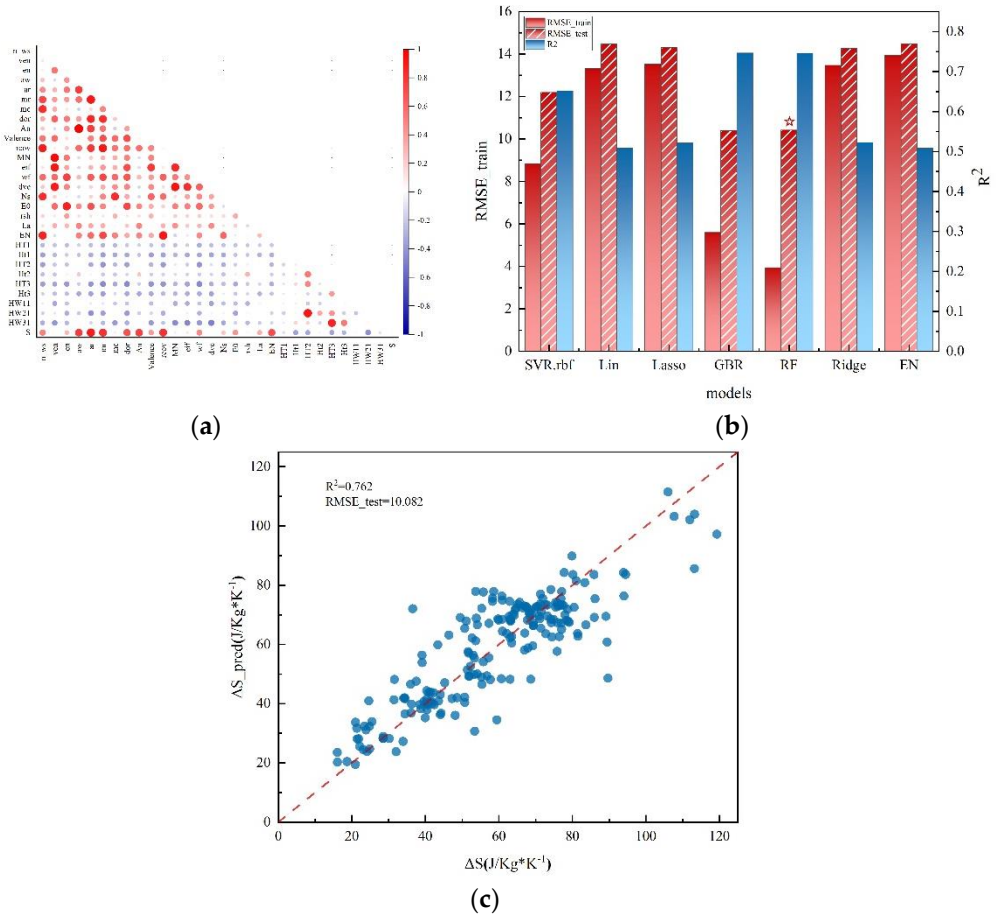

Figure 2. Characteristics and model selection for physical and chemical characteristics combined with heat treatment process characteristics model (M2). (a) Pearson correlation matrix for the initial 29 descriptors. (b) Model selection based on leave-out cross-validation. (c) The prediction performance of M2.

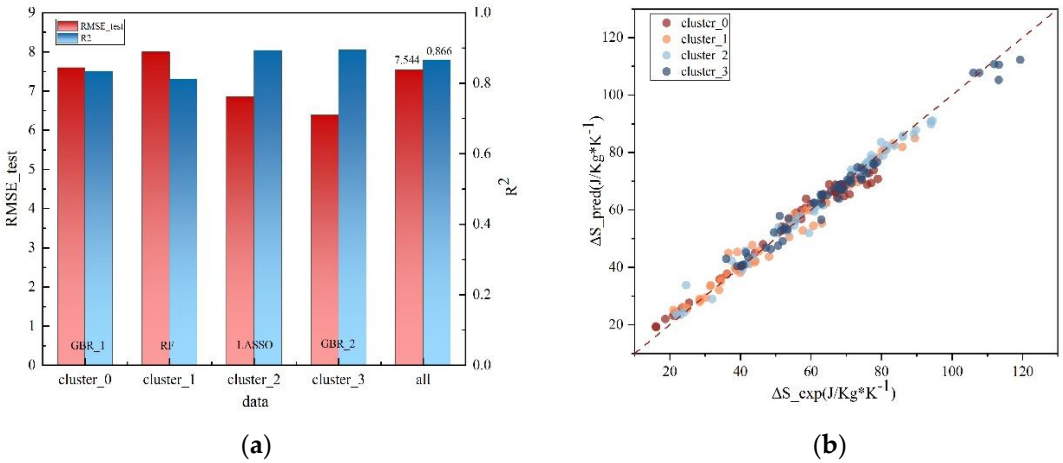

Figure 3. Results of K-Means clustering model M3. (a) The best models for each clustering subset, and their R2 and RMSE. (b) The prediction performance of M3.

 doi:10.20944/preprints202408.1415.v1

Table 2. The best models for four clustering subsets, their R^2 and RMSE, and comprehensive results for K-Means clustering model M3.

Clusters	Models	R^2	RMSE_train	RMSE_test
cluster_0	GBR_1	0.834	3.178	7.598
cluster_1	RF	0.812	2.899	7.998
cluster_2	Lasso	0.893	5.678	6.861
cluster_3	GBR_2	0.895	1.859	6.393
all		**0.866**		**7.544**

In physics, entropy is commonly understood as the degree of disorder within a system. The entropy ΔS of an alloy is influenced by various factors, as expressed in Eq. (3)[37]

$$\Delta S = \Delta S_{vib.} + \Delta S_{el.} + \Delta S_{mag.} + \Delta S_{dis.} + \Delta S_{conf.} \quad (3)$$

where ΔS_{vib} represents the entropy change caused by atomic vibration, ΔS_{ele} represents the entropy change caused by electron, ΔS_{mag} represents the entropy change caused by magnetic properties of material, and ΔS_{conf} represents the entropy change caused by structural alteration.

Given that the NiTi-based SMAs in this study is non-magnetic, we have $\Delta S_{mag} = 0$. Additionally, since the martensitic phase change is a non-diffusive phase transformation, $\Delta S_{conf} = 0$.

While the highest-scoring four features for M1 differ from those in M2 and M3, all three models encapsulate information about the alloy at the electronic, atomic, and lattice levels. Electronegativity (en), for instance, denotes the atom's ability to attract electrons. Electronegativity, valence electron number (ven) and covalent radius (rcov) exhibit periodic trends, with main group elements of the same period witnessing an increase in en and a decrease in rcov as ven increases. In the same main group, as periodicity increases, en decreases, while rcov increases.

Average electron concentration (aven) and the s orbital electron filling number (NsValence) are related to the element's position in the periodic table. In addition, the compression modulus has a relationship, as per Eq. (4), with the elastic modulus.

$$K = \frac{E}{3(1-2v)} \quad (4)$$

where E represents the modulus of elasticity, K represents the modulus of compression, and v is the Poisson's ratio of the alloys, signifying lateral shrinkage capacity. The elastic modulus indicates the ease of atoms leaving their equilibrium position, characterizing the strength of interatomic bonding. Crystals with strong interatomic bonding, such as covalently bonded crystals like diamond, have a high elastic modulus.

Features like aven, en, Nsvalence, compression modulus, rcov and ven all directly or indirectly reflect information about ΔS_{ele} at the electronic level. Additionally, the effect of atomic mass (am) on vibrational entropy shows that the larger the atomic mass in a compound leads to less violent vibration under the same conditions, resulting in a smaller ΔS_{vib}. Despite the non-identical feature sets, both provide a reasonable explanation for the phase transformation entropy change ΔS in terms of electron, atomic and lattice aspects.

3.2. Validation

To evaluate the machine learning model, we not only assessed its performance on the training data but also prioritized examining its extrapolation ability. For this purpose, we collected 5 new data points that were not present in the training dataset to assess the extrapolation ability of M1. Similarly, 30 new data points were gathered for validating M2 and M3. For the M3 segmented fit model, a preliminary step involved assigning each data point to the relevant cluster before entering it into the prediction model. Here, we computed the clustering center of each cluster model Ci (i = 1,2,3,4), where Ci represented a sequence of 13 × 1. The next step involved calculating the Euclidean distance (Dis_{ij} =

 doi:10.20944/preprints202408.1415.v1

$\sqrt{(D_j - C_i)^2}$) of each data point ($D1$, …, Dj, …, $D30$) to the four clustering centers separately. We considered a data D_j to belong to cluster C_i when the distance Dis_{ij} was minimal.

The validation results for the new data of M1, M2, and M3 are shown in Figure 4 (a), (b) and (c), respectively, with average error of 13.84%,13.51%, 2.91%. Notably, the incorporation of heat treatment process features led to a significant reduction in the extrapolation error for the model. Moreover, the segmented fitting model proved to be even more effective in minimizing extrapolation error, providing clear evidence of the improved extrapolation ability of the model.

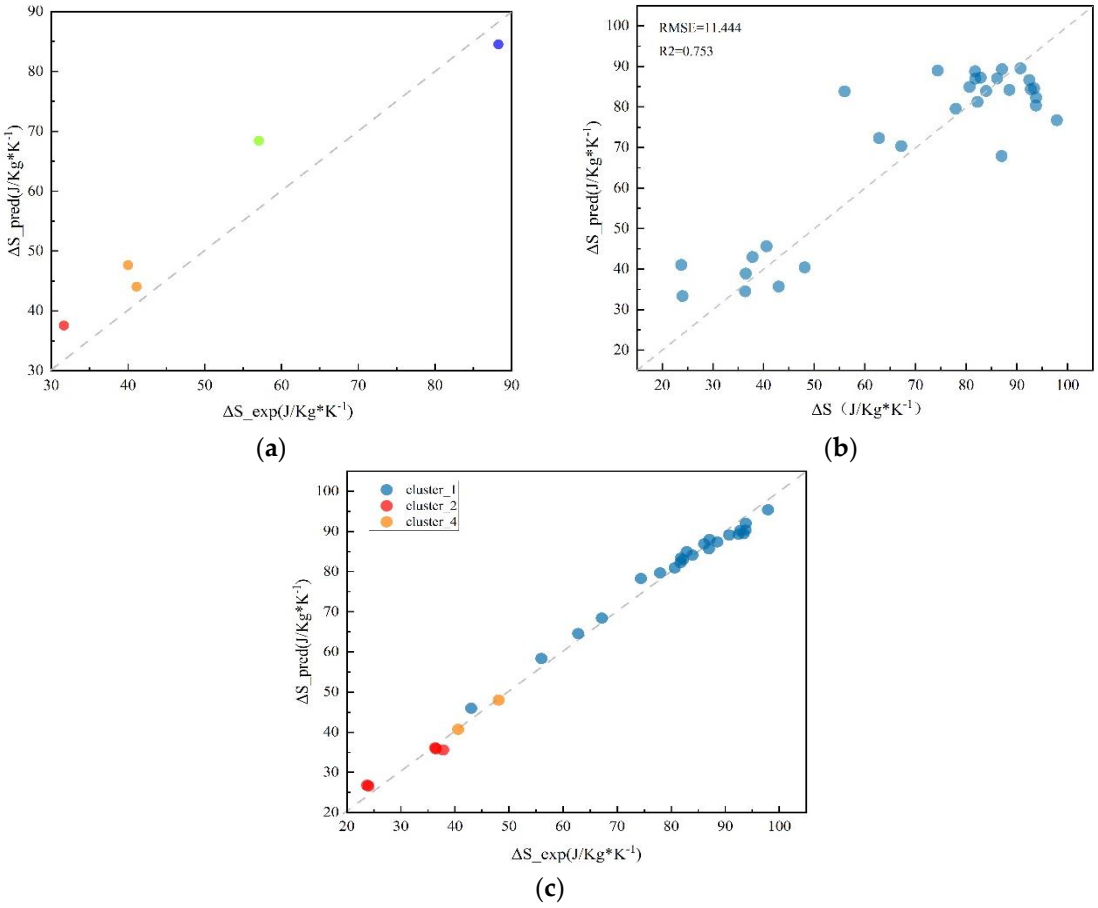

Figure 4. Fitting results of M1 (a), M2 (b), M3 (c) on new validation data.

3.3. Materials Design toward High Elastocaloric SMA

In pursuit of NiTi-based SMAs with heightened elastocaloric effects, we apply machine learning models M1, M2, and M3 to an adaptive design approach. The exploration encompassed the search space for binary and ternary SMAs, comprising 51 NiTi-based SMA system. The third group of elemental predictions excluded gases, radioactive elements, toxic elements, alkali metal elements, halogens with lanthanum representing the lanthanide system. The elemental concentration ranges are detailed in Table 3, with a step of 0.5 at.% for each element. In addition, we defined the search space for heat treatment process parameters for M2 and M3, as outlined in Table 4. Notably, we constrained the parameters for homogenization heat treatment to 1323 K, 4320 mins, and water quenching, optimizing for computational efficiency and cost-effectiveness. Additionally, for solution and aging heat treatment, we introduced values not only from the Table 4 but also used the heat treatment process parameters with temperature of 293 K, time of 0 mins, and cooling method of air cooling to complete the missing process parameters.

 doi:10.20944/preprints202408.1415.v1

Table 3. Types and concentration ranges of third group elements in NiTi-based SMA for M1, M2 and M3.

Element	Range
Ni	1-Ti-third element
Ti	45 at. % ~ 60 at. %
Cu、V、Mn、Fe、Hf、Pd	0 ~ 30 at. %
Other	0 ~ 10 at. %

Table 4. Range and step of 9 heat treatment process parameters for M2 and M3. HT1, Ht1, CQ1 indicate the temperature, time and ways of cooling of Homogenization, HT2, Ht2, CQ2 indicate the parameter of Solution, and HT3, Ht3, CQ3 indicates the parameter of Aging.

Heat treatment	Range	Step
HT_{I}	1323 K	/
Ht_{I}	4320 mins	/
W_{I}	WQ	/
HT_{II}	323 K~ 1223 K	300K
Ht_{II}	60 mins~ 720 mins	60mins
W_{II}	WQ/AC	/
HT_{III}	323 K~ 1223 K	300K
Ht_{III}	30 mins ~ 420 mins	60mins
W_{III}	WQ/AC	/
HT_{I}	1323 K	/
Ht_{I}	4320 mins	/
W_{I}	WQ	/

For M1, candidate alloys within the search space were selected based on the maximum expected improvement $EI\ (\mu,\ \sigma)$[32] as defined by Eq.6. Active learning, as shown in Figure 1 (e), led to the discovery of 9 new NiTi-based SMAs with ΔS>90 J/Kg·K-1 in four iterations, as detailed in Supplementary Table 3.

$$EI(\mu,\sigma) = \sigma[\phi(z) + z\Phi(z)] \tag{6}$$

Here, $z = (\mu - \mu^{*})/\sigma$, where $\mu^{*}$is the maximum value in the training set, μ is the mean value of the predicted data, and σ is the standard deviation of $\phi(z)$ and$\Phi(z)$, representing the standard normal density and distribution function, respectively. Similarly, active learning was also employed for M2 and M3, and the initial results closely mirrored those of M1. Due to timeliness consideration, the subsequent experimental results will be published separately.

4. Summary

In this work, we harnessed the power of machine learning to expedite the discovery of novel NiTi-based SMAs with high elastocaloric effect, as well as optimizing the corresponding heat treatment processes. Our approach involved constructing machine learning prediction models, specifically the SVR.rbf model (M1), which utilized four physicochemical features (aven, en, Nsvalence, Modulus compression). M1 exhibited superior performance with an R2=0.860 and an RMSE=8.472. Leveraging an active learning strategy through four iterations, we identified 9 new NiTi-based SMAs with phase transformation entropy change ΔS>90 J/Kg·K-1, surpassing the majority of alloys in the original dataset.

For practical engineering consideration, we extended our efforts to establish the RF model (M2) for ΔS, incorporating 4 physicochemical features (aw, dor, rcov, Valence). To enhance interpretability, we performed feature selection and introduced 9 heat treatment process features as inputs.

 doi:10.20944/preprints202408.1415.v1

Consequently, the model's performance slightly diminished, yielding an R2= 0.746 and an RMSE=10.422. To address this, we implemented K-Means clustering to formulate the model M3, achieving improved prediction performance with an R2= 0.866 and an RMSE= 7.544. Active learning was similarly applied to guide alloys design, yielding results in line with M1 in the two iterations.

It's noteworthy that the selected features in M1, M2 and M3 encompass information about NiTi-based SMA at the electronic, atomic and lattice levels. Furthermore, M2 and M3 incorporate the influence of the heat treatment process. Depending on specific requirements, practitioners can choose different models to assist their decision-making processes.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Author Contributions: Data collection, modeling, validation, formal analysis, investigation, and writing—original draft preparation, Y.G. and Y.H.; methodology and validation, X.Z. and Y.L.; conceptualization, writing—review and editing, and supervision, H.H.; resources, project administration, funding acquisition, H.H. and Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Guangdong Major Project of Basic and Applied Basic Research, grant number No.2019B030302011.

Data Availability Statement: The raw/processed data required to reproduce these findings can be obtained by contacting the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Secretariat O. The Montreal protocol on substances that deplete the ozone layer. United Nations Environment Programme. City: Nairobi, Kenya, 2000:54.
2. Coulomb D, Dupont J-L, Pichard A. The Role of Refrigeration in the Global Economy - 29 Informatory Note on Refrigeration Technologies. City: Paris, 2015.
3. Franco V, Blázquez J S, Ingale B, et al. The magnetocaloric effect and magnetic refrigeration near room temperature: materials and models. Annual Review of Materials Research, 2012, 42: 305-342.
4. Valant M. Electrocaloric materials for future solid-state refrigeration technologies. Progress in Materials Science, 2012, 57(6): 980-1009.
5. Qian S, Geng Y, Wang Y, et al. A review of elastocaloric cooling: Materials, cycles and system integrations. International journal of refrigeration, 2016, 64: 1-19.
6. Goetzler W, Shandross R, Young J, et al. Energy savings potential and RD&D opportunities for commercial building HVAC systems. City: Navigant Consulting, Burlington, MA (United States), 2017.
7. Cui J, Wu Y, Muehlbauer J, et al. Demonstration of high efficiency elastocaloric cooling with large ΔT using NiTi wires. Applied Physics Letters, 2012, 101(7): 073904.
8. Ossmer H, Miyazaki S, Kohl M. Elastocaloric heat pumping using a shape memory alloy foil device// Elastocaloric heat pumping using a shape memory alloy foil device. 2015 Transducers-2015 18th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS), Anchorage, AK, USA :IEEE, : 726-729.
9. Schmidt M, Kirsch S-M, Seelecke S, et al. Elastocaloric cooling: From fundamental thermodynamics to solid state air conditioning. 2016, 22(5): 475-488.
10. Tušek J, Engelbrecht K, Millán-Solsona R, et al. The elastocaloric effect: a way to cool efficiently. Advanced Energy Materials, 2015, 5(13): 1500361.
11. Otsuka K, Ren X J P i m s. Physical metallurgy of Ti–Ni-based shape memory alloys. Progress in materials science, 2005, 50(5): 511-678.
12. Frenzel J, George E P, Dlouhy A, et al. Influence of Ni on martensitic phase transformations in NiTi shape memory alloys. Acta Materialia, 2010, 58(9): 3444-3458.
13. Ma J, Karaman I, Noebe R D. High temperature shape memory alloys. International Materials Reviews, 2010, 55(5): 257-315.
14. Zhou Y, Xue D, Ding X, et al. Strain glass in doped $Ti_{50}(Ni_{50-x}D_x)$ (D=Co, Cr, Mn) alloys: Implication for the generality of strain glass in defect-containing ferroelastic systems. Acta Materialia, 2010, 58(16): 5433-5442.
15. Ossmer H, Miyazaki S, Kohl M. The Elastocaloric Effect in TiNi-based Foils. Materials Today: Proceedings, 2015, 2: S971-S974.
16. Schmidt M, Schütze A, Seelecke S. Scientific test setup for investigation of shape memory alloy based elastocaloric cooling processes. International Journal of Refrigeration, 2015, 54: 88-97.

 doi:10.20944/preprints202408.1415.v1

17. Chen H, Xiao F, Li Z, et al. Elastocaloric effect with a broad temperature window and low energy loss in a nanograin Ti-44Ni-5Cu-1Al Ti-44Ni-5Cu-1Al (at·%) shape memory alloy. Physical Review Materials, 2021, 5(1): 015201.
18. Xu B, Wang C, Wang Q, et al. Enhancing elastocaloric effect of NiTi alloy by concentration-gradient engineering. International Journal of Mechanical Sciences, 2023, 246: 108140.
19. Yeung K W K, Cheung K M C, Lu W W, et al. Optimization of thermal treatment parameters to alter austenitic phase transition temperature of NiTi alloy for medical implant. Materials Science and Engineering: A, 2004, 383(2): 213-218.
20. Deng Z, Huang K, Yin H, et al. Temperature-dependent mechanical properties and elastocaloric effects of multiphase nanocrystalline NiTi alloys. Journal of Alloys and Compounds, 2023, 938: 168547.
21. Liu S, Kappes B B, Amin-ahmadi B, et al. Physics-informed machine learning for composition–process–property design: Shape memory alloy demonstration. Applied Materials Today, 2021, 22: 100898.
22. Malinov S, Sha W, McKeown J J. Modelling the correlation between processing parameters and properties in titanium alloys using artificial neural network. Computational Materials Science, 2001, 21(3): 375-394.
23. 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.
24. Liu P, Huang H, Antonov S, et al. Machine learning assisted design of γ′-strengthened Co-base superalloys with multi-performance optimization. npj Computational Materials, 2020, 6(1): 62.
25. Ohno H. Training data augmentation: An empirical study using generative adversarial net-based approach with normalizing flow models for materials informatics. Applied Soft Computing, 2020, 86: 105932.
26. Chanda B, Jana P P, Das J. A tool to predict the evolution of phase and Young's modulus in high entropy alloys using artificial neural network. Computational Materials Science, 2021, 197: 110619.
27. Lee J, Asahi R. Transfer learning for materials informatics using crystal graph convolutional neural network. Computational Materials Science, 2021, 190: 110314.
28. Ding Y, Liu C, Zhu H, et al. A supervised data augmentation strategy based on random combinations of key features. Information Sciences, 2023, 632: 678-697.
29. He S, Wang Y, Zhang Z, et al. Interpretable machine learning workflow for evaluation of the transformation temperatures of TiZrHfNiCoCu high entropy shape memory alloys. Materials & Design, 2023, 225: 111513.
30. Yang Z, Li S, Li S, et al. A two-step data augmentation method based on generative adversarial network for hardness prediction of high entropy alloy. Computational Materials Science, 2023, 220: 112064.
31. Zhang Y-F, Ren W, Wang W-L, et al. Interpretable hardness prediction of high-entropy alloys through ensemble learning. Journal of Alloys and Compounds, 2023, 945: 169329.
32. Xue D, Balachandran P V, Hogden J, et al. Accelerated search for materials with targeted properties by adaptive design. Nat Commun, 2016, 7: 11241.
33. Xue D, Xue D, Yuan R, et al. An informatics approach to transformation temperatures of NiTi-based shape memory alloys. Acta Materialia, 2017, 125: 532-541.
34. Maharana K, Mondal S, Nemade B. A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings, 2022, 3(1): 91-99.
35. Gareth J, Daniela W, Trevor H, et al. An introduction to statistical learning: with applications in R. City: Spinger, 2013.
36. Schapire R E J N e, classification. The boosting approach to machine learning: An overview. Nonlinear estimation and classification, 2003, 171: 149-171.
37. Khalil-Allafi J, Amin-Ahmadi B. The effect of chemical composition on enthalpy and entropy changes of martensitic transformations in binary NiTi shape memory alloys. Journal of Alloys and Compounds, 2009, 487(1-2): 363-366.
38. Wang T, Zhang K, Thé J, et al. Accurate prediction of band gap of materials using stacking machine learning model. Computational Materials Science, 2022, 201: 110899.
39. Liu Y, Wu J, Wang Z, et al. Predicting creep rupture life of Ni-based single crystal superalloys using divide-and-conquer approach based machine learning. Acta Materialia, 2020, 195: 454-467.
40. Zhang S, Chen D, Liu S, et al. Aluminum alloy microstructural segmentation method based on simple noniterative clustering and adaptive density-based spatial clustering of applications with noise. Journal of Electronic Imaging, 2019, 28(3): 033035.
41. Malinov S, Sha W, McKeown J J. Modelling the correlation between processing parameters and properties in titanium alloys using artificial neural network. Computational materials science, 2001, 21(3): 375-394.
42. Waber J, Cromer D T. Orbital radii of atoms and ions. The Journal of Chemical Physics, 1965, 42(12): 4116-4123.
43. Pauling L. The nature of the chemical bond. IV. The energy of single bonds and the relative electronegativity of atoms. Journal of the American Chemical Society, 1932, 54(9): 3570-3582.
44. Rabe K M, Phillips J C, Villars P, et al. Global multinary structural chemistry of stable quasicrystals, high-TC ferroelectrics, and high-Tc superconductors. Physical Review B, 1992, 45(14): 7650-7676.

 doi:10.20944/preprints202408.1415.v1

45. Clementi E, Raimondi. Atomic screening constants from SCF functions. The Journal of Chemical Physics, 1963, 38(11): 2686-2689.
46. Pettifor D G, Pettifor D. Bonding and structure of molecules and solids. City: Clarendon Press Oxford, 1995.