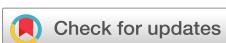


RESEARCH ARTICLE | JULY 05 2023

## Machine learning guided optimal composition selection of niobium alloys for high temperature applications

Trupti Mohanty  ; K. S. Ravi Chandran  ; Taylor D. Sparks  



*APL Mach. Learn.* 1, 036102 (2023)  
<https://doi.org/10.1063/5.0129528>



### Articles You May Be Interested In

Predicting shear modulus property using materials informatics

*AIP Conf. Proc.* (August 2024)

Machine learning-based inverse design for single-phase high entropy alloys

*APL Mater.* (October 2022)

Predicting composition and bulk modulus property linkage using materials informatics

*AIP Conf. Proc.* (August 2024)

# Machine learning guided optimal composition selection of niobium alloys for high temperature applications

Cite as: APL Mach. Learn. 1, 036102 (2023); doi: 10.1063/5.0129528

Submitted: 5 October 2022 • Accepted: 5 June 2023 •

Published Online: 5 July 2023



View Online



Export Citation



CrossMark

Trupti Mohanty, K. S. Ravi Chandran, and Taylor D. Sparks<sup>a)</sup>

## AFFILIATIONS

Department of Materials Science and Engineering, University of Utah, Salt Lake City, Utah 84112, USA

<sup>a)</sup>Author to whom correspondence should be addressed: [sparks@eng.utah.edu](mailto:sparks@eng.utah.edu)

## ABSTRACT

Nickel- and cobalt-based superalloys are commonly used as turbine materials for high-temperature applications. However, their maximum operating temperature is limited to about 1100 °C. Therefore, to improve turbine efficiency, current research is focused on designing materials that can withstand higher temperatures. Niobium-based alloys can be considered as promising candidates because of their exceptional properties at elevated temperatures. The conventional approach to alloy design relies on phase diagrams and structure–property data of limited alloys and extrapolates this information into unexplored compositional space. In this work, we harness machine learning and provide an efficient design strategy for finding promising niobium-based alloy compositions with high yield and ultimate tensile strength. Unlike standard composition-based features, we use domain knowledge-based custom features and achieve higher prediction accuracy. We apply Bayesian optimization to screen out novel Nb-based quaternary and quinary alloy compositions and find these compositions have superior predicted strength over a range of temperatures. We develop a detailed design flow and include Python programming code, which could be helpful for accelerating alloy design in a limited alloy data regime.

© 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0129528>

## INTRODUCTION

Gas turbines have wide applications in power generation and aerospace industries, which operate at very high temperatures. The efficiency of a turbine can be increased by increasing the operating temperature,<sup>1</sup> which, in turn, requires the use of alloys with higher strength levels relative to those currently in use. At present, Ni-based superalloys are used in turbines that can operate up to 1100 °C. Therefore, there is a need to design turbine materials that can withstand temperatures >1100 °C to realize further improvement in turbine efficiency. Refractory alloys based on Nb, W, and Mo have the potential for high-temperature applications because of their exceptional properties at elevated temperatures. Among these, niobium is a promising candidate as a turbine material due to its high melting point, low density, high room temperature ductility, and low ductile to brittle temperature.<sup>2</sup> High-temperature, high-strength Nb-based alloys can be realized by adding suitable alloying elements that can provide solid solution, dispersion, and multi-phase strengthening, as is well-known in the physical metallurgy of

refractory alloys.<sup>3,4</sup> However, these alloying elements yield a huge compositional design space. Suppose, for example, we have 14 alloying elements, excluding the base element niobium, and each alloying element's composition varies from 1 to 20 wt. % with a step size of 1 wt. %. In that case, the total number of possible alloy combinations can be expressed as  ${}_{n-1}^{14}C \times 20^{n-1}$ , where  $n$  represents the number of elements in the alloy system. For example, for a quinary alloy system,  $n = 5$ . With this, there are  $\sim 3 \times 10^6$  quaternary and  $\sim 1.6 \times 10^8$  quinary possible alloy combinations. Clearly, tools are needed to efficiently evaluate this immense design space. To this end, data-driven methods can complement the knowledge-based approaches, such as computational simulations, as well as thermodynamic and empirical modeling, currently used for alloy development.

In recent years, machine learning (ML) has emerged as an efficient tool for speeding up materials research. This could be possible due to the availability of large experimental and computational datasets, improvement in algorithms, and growth in computer processing power. ML can predict material properties, accelerate the discovery of new materials, narrow down relevant composition

domains, and design efficient experimental methods based on optimization algorithms.<sup>5–8</sup> In fact, accurate predictions of material properties have been shown using composition alone. For example, Goodall and Lee have developed a ML approach referred to as Roost (Representation Learning from Stoichiometry), which takes only the material composition as input in the form of a dense weighted graph and successfully predicts material properties.<sup>6</sup> Similarly, CrabNet (Compositionally restricted attention-based network) is a machine learning network based on the transformer attention mechanism that takes elemental chemistry as input and predicts the material properties using composition only. CrabNet performs exceptionally well in almost all 28 benchmark datasets,<sup>7</sup> even when compared with algorithms that use both composition and structure as descriptors. Not only in predicting material properties, but ML has also been successfully applied to new material discoveries.<sup>8</sup> Mansouri Tehrani *et al.* discovered two new promising superhard materials, rhenium tungsten carbide, and molybdenum tungsten borocarbide, by training a support vector regressor (SVR) model on computational density functional theory (DFT) data followed by screening from an extensive known crystal structure database.<sup>8</sup> The reported hardness value for each of these newly discovered compounds exceeds 40 GPa at low load (0.49 N).

ML has been extensively applied to alloy design,<sup>9</sup> such as in the design of high entropy alloys, titanium alloys, copper alloys, shape memory alloys, and bulk metallic glass alloys. For the development of copper alloys with targeted properties, such as ultimate tensile strength (UTS) and electrical conductivity, property-oriented machine learning-based design strategies have been used.<sup>10</sup> In a similar approach, Nazarhari and Canadinc used a multilayer feed-forward neural network to design the optimum composition of Ni-Ti superalloy for dental application.<sup>11</sup> The optimum composition of 51.5 at. % Ni with balance Ti showed the lowest amount of

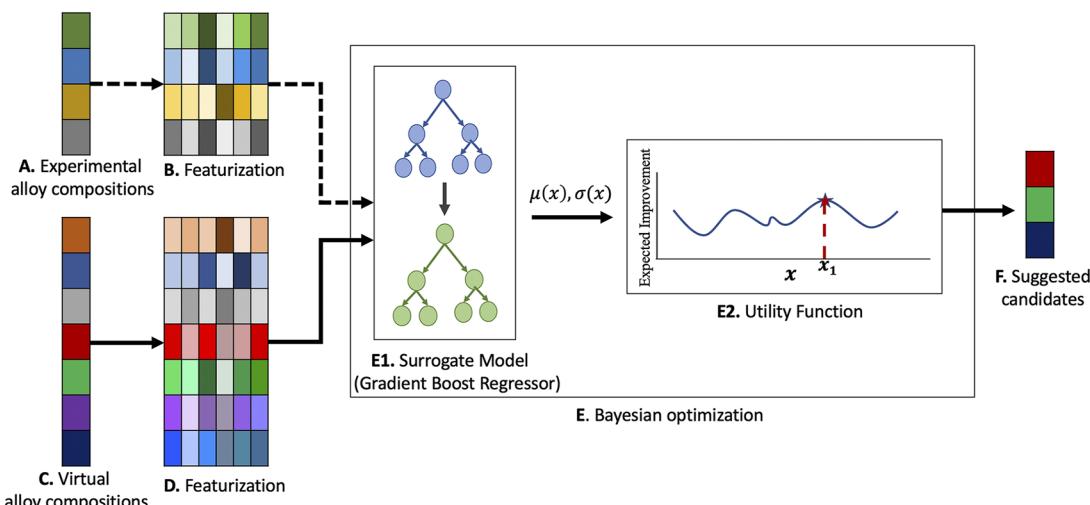
nickel ion release into the dental cavity. Recently, Wen *et al.*<sup>12</sup> have successfully applied Bayesian optimization (BO) based active learning in AlCoCrCuFeNi high entropy alloy system and discovered new alloy compositions having hardness value 10% higher than the best value in their training dataset.

Although ML has been proven successful in predicting material properties in various studies, applying the common ML strategies in the case of Nb is challenging because of inadequate available experimental data. Moreover, the gas turbines are repeatedly cycled from hot to cold during their operation; thus, it is required to discover novel materials that likely pose higher strength over a broader range of temperatures (room temperature – 1300 °C). In this current study, we addressed these challenges by incorporating domain knowledge based custom features along with temperature to train the ML model. We adopted the BO framework for screening the vast composition space and suggested promising Nb alloy candidates having higher predicted strength over the desired temperature range. This computational approach will considerably reduce the experimental efforts, enabling accelerated and more economic development of niobium alloys.

## METHODS

### Machine learning strategy

A schematic workflow for the machine learning guided alloy design strategy has been provided in Fig. 1. We carried out featurization to represent each experimental alloy composition in a numerical vector format using domain knowledge-based customized material descriptors (Figs. 1A and 1B). We constructed a list of virtual quaternary and quinary alloy compositions and performed featurization in



**FIG. 1.** Schematic workflow for the machine learning guided alloy design using Bayesian optimization (BO). A. Existing experimental Nb alloy compositions. B. Featurization: Each row represents an alloy composition in numerical vector format. C. Virtual alloy compositional space. D. Featurization for virtual alloy candidates: Each row represents an alloy composition in numerical vector format. E. Bayesian optimization module. E1. Surrogate model: Gradient Boost Regressor (GBR) model trained on the available dataset. E2. Utility function: Uses mean and standard deviation and suggest suitable alloy candidates. F. Suggested alloy candidates. The dotted lines indicate the data flow to train the model, and the solid lines represent the data flow for the prediction of mechanical properties and suggestion of alloy candidates.

a similar way (Figs. 1C and 1D). To screen out suitable alloy candidates from this virtual candidate search space, we use Bayesian optimization (Fig. 1E). Bayesian optimization (BO) is a sequential design strategy that utilizes a surrogate model (Fig. 1E1) and a utility function (Fig. 1E2) for finding the optimal alloy candidates. In each iteration of BO, the surrogate model is trained on the experimental alloy data. For each virtual alloy candidate  $x$ , the trained model estimates the mean strength  $\mu(x)$  and standard deviation  $\sigma(x)$  as a measure of uncertainty in the prediction. The utility function uses these estimated means and standard deviations to suggest the best possible candidate  $x_1$  to carry out the next experiment. In this work, we have used Gradient Boost Regressor (GBR) as a surrogate model and expected improvement (EI) as utility function. The selection of surrogate model and the quantification of means and standard deviations are detailed in subsequent sections.

### Data collection and featurization

We compiled the experimental mechanical property data, such as ultimate tensile strength (UTS) and yield strength (YS) data, for known niobium alloys.<sup>13</sup> This dataset consists of 18 unique Nb alloys where mechanical strength is reported at varying temperatures ranging from 24 to 1871 °C (Figs. S1 and S2, supplementary material). This leads to a total of 140 yield strength and 144 ultimate tensile strength data points. Out of total 144 data points, 140 contain both ultimate tensile strength and yield strength values. The dataset does not have the yield strength value of Nb<sub>72.8</sub>C<sub>0.7</sub>Mo<sub>8.8</sub>Ti<sub>17.7</sub> at 24 and 1260 °C and yield strength value of Nb<sub>64.4</sub>W<sub>10.2</sub>Ti<sub>19.5</sub>Mo<sub>5.8</sub> at 1260 and 1427 °C.

The goal was to train a machine learning model to accurately predict the targeted mechanical property based on the composition/feature–property relationship and evaluate its performance on unseen test data. To use the machine learning algorithms and to evaluate their performance, we split the data into train and test sets with 80:20 train-test split ratio. We ensured that data splitting was carried out not only to prevent data leakage from duplicate data points but also to ensure that the whole alloy families were grouped into either train or test to maximize model generalizability.

In this work, we compared three different featurization schemes: elemental atom percentage, extensive composition-based featurization, and domain knowledge-based custom feature set. Elemental atom percentage is one of the simplest approaches for featurization in materials informatics,<sup>14</sup> which uses the atom percentage values of the alloying elements to construct a feature vector. While this approach can work well with large datasets,<sup>15</sup> in the case of a limited dataset (e.g., only 18 unique alloys), ML models will likely fail to capture the underlying chemistry between the alloying elements and the property. Indeed, we have shown this to be true in previous materials informatics studies.<sup>16</sup> Accordingly, the trained model will also be unlikely to perform well while screening the optimum alloy composition from an unexplored design space with unseen alloying elements that are not present in the training dataset. An alternative to simple elemental atom percentage would be the construction of an extensive composition-based feature vector (CBFV) using an elemental descriptor set, such as Oliynyk.<sup>17</sup> The composition-based feature vector for each alloy was derived using the CBFV python library.<sup>17</sup> From this library, we have used the Oliynyk elemental descriptor, which consists of 44 elemental properties, such as atomic

number, atomic weight, period, group, etc. (complete list available in GitHub repository). CBFV featurization scheme uses these 44 elemental properties from the Oliynyk elemental descriptors and then provides descriptive statistics, such as weighted average, range, and variance, for alloy compositions to generate a final descriptor that is a 132-dimensional vector. The potential downside of using such a large feature vector with limited data is likely to overfit. Finally, as a compromise between simple elemental atom percentage and overly extensive Oliynyk CBFV, we consider physical metallurgy guided custom feature set of reduced dimensions (15 features) consisting of elemental descriptors closely related to yield strength and ultimate tensile strength. These elemental descriptors are atomic radius, Pauling electronegativity, cohesive energy, number of valence electrons, bulk modulus, elastic modulus, shear modulus, rate of change of shear modulus with temperature, melting temperature, maximum solid solubility limit of the elements in niobium, average valence bond strength, bulk electron concentration, lattice constant, Engel's net bonding valence electrons e/a ratio, and the mechanical properties test temperature. These elemental descriptor data are collected from the literature<sup>18–20</sup> and compiled. The feature vector of each alloy was created using a weighted average in the form  $\sum_{i=1}^n c_i f_i$ , where  $c_i$  represents atom fraction of each element  $i$  present in the alloy and  $f_i$  represents corresponding elemental descriptor.

Working examples of this different featurization schema are provided in our GitHub repository, and the results are shown in Fig. S3, supplementary material. We created the feature matrix for the train and test dataset. Following the best practices,<sup>17</sup> we normalized the feature matrix of the training dataset to zero mean and unit variance. We used the mean and variance of the train dataset feature matrix to transform the test dataset feature matrix. The normalized training dataset is used for training the ML model.

### Machine learning regression models and uncertainty estimation

We have used several well-known supervised regressor machine learning algorithms such as Gradient Boost regressor (GBR), Random Forest (RF) regressor, Support Vector Regressor (SVR), K-nearest neighbors (KNN) regressor, and Multi-Layer Perceptron (MLP) for this study. These models are trained using the training data, and their performances are evaluated using the test data. The hyperparameters of the models were tuned using grid search with five-fold cross-validation. In five-fold cross-validation, the training data were divided into five smaller datasets known as folds, and the model was trained on four of the folds as training data and validated on the remaining part of the data. The hyperparameters were tuned based on the average value of the performance scores computed in five different loops. Here, we have used the python library scikit-learn GroupKFold function to ensure that the same alloy composition does not repeat in different folds to avoid information leakage within folds. Once the hyperparameters of the model were tuned, performance evaluation of the model was carried out using the unseen test data. Mean absolute error (MAE), root mean square error (RMSE), and the correlation coefficient ( $R^2$ ) were used as metrics to evaluate the performance of the machine learning models. In this study, models not only predicted the targeted property but also quantified the uncertainty of the predictions. The uncertainty in the predictions of the best-performing machine

learning model was estimated by using the bootstrap technique. In the bootstrap technique, multiple training datasets are generated by random sampling from the original training dataset with replacement while keeping the size of each sampled dataset the same as our original training dataset. After that, multiple models are developed using these sampled training datasets while keeping the model's hyperparameters constant. We have created 1000 such bootstrapped training datasets and trained 1000 machine learning models. Thus, for each test candidate, 1000 predictions were obtained from which the uncertainties, such as the mean and standard deviation of the prediction, are estimated.

### Generation of virtual candidate search space

Virtual quaternary alloy compositions were created of the form  $\text{Nb}_a\text{X}_x\text{Y}_y\text{Z}_z$ , where X, Y, and Z represents any of the 14 alloying elements B, C, Co, Cr, Hf, Mo, N, Re, Si, Ta, Ti, V, W, Zr, and  $x, y$ , and  $z$  represent their corresponding weight percent value and  $a = [100 - (x + y + z)]$  represents the weight percent of Nb. These 14 alloying elements were selected based on domain knowledge of the promising candidates to provide the solid solution, precipitation, and dispersion strengthening. The concentration granularity of the alloying elements was 1 wt %. The concentration of B, C, and N varied from 1 to 5 wt %, and for the rest of the elements, concentration varied from 1 to 20 wt %. This combination yielded 1 666 625 ( $\sim 1.6 \times 10^6$ ) unique quaternary alloy compositions. With an additional constraint of niobium weight percentage  $\geq 50$ , the search space yielded 1 630 325 alloy candidates. With a similar approach, quinary alloy compositions of the form  $\text{Nb}_a\text{W}_w\text{X}_x\text{Y}_y\text{Z}_z$  have been generated, where W, X, Y, and Z represent any of the 14 alloying elements and  $w, x, y$ , and  $z$  represent their corresponding weight percent value and  $a = [100 - (w + x + y + z)]$  represents the weight percent of Nb. A total of 74 277 500 ( $\sim 74 \times 10^6$ ) quinary alloy compositions were created. After adding a constraint of niobium weight percentage  $\geq 50$ , the search space yielded 60 556 925 ( $\sim 60 \times 10^6$ ) alloy candidates. The next step was to screen out promising alloy compositions from this vast virtual candidate search space.

### Bayesian optimization for screening promising candidates

As the ML model is built using a limited number of training data, the candidates' selection using the model might be limited to local search. Therefore, we surmised that Bayesian optimization (BO) would likely give a better result as this optimization technique considers the prediction uncertainty and balances between local and global search. BO works on a surrogate model and evaluates a utility function.<sup>21</sup> The utility function uses the mean and the standard deviation of the predictions estimated by the surrogate model. The utility function encodes a trade-off between exploitation (candidate searching at points with high mean) and exploration (candidate searching at points with high uncertainty). Based on the output of the utility function, BO suggests alloy compositions to carry out the next experiment. As BO is an active learning based iterative approach for global optimization, these suggested samples can be evaluated and fed back to the training data for further improvement of the performance of the surrogate model during repeated process iterations. The iterations can be stopped if the strength value of the selected

candidate is more than the highest strength value observed in the training dataset.

In this study, we have used GBR as the surrogate model and expected improvement (EI) as a utility function. The utility function is defined as  $EI(x) = \sigma(x)[z\Phi(z) + \phi(z)]$ , where  $EI(x)$  represents the expected improvement value for each alloy candidate,  $z = [\mu(x) - f(x^+) - \epsilon]/\sigma(x)$ ,  $\mu$  and  $\sigma$  are mean prediction and standard deviation, respectively,  $f(x^+)$  is the maximum value of the target material property observed in the training dataset.  $\Phi$  represents the cumulative distribution function and  $\phi$  is the probability distribution function assuming the target property values follows the normal distribution. The term  $\epsilon$  regulates the amount of exploration; a higher value of  $\epsilon$  encourages more exploration. The candidate corresponds to largest EI value is the most promising one. We evaluated the EI values for all alloy compositions at room temperature because (a) the maximum observed strength values of the existing alloys are at room temperature, and (b) we hypothesize that the alloys having high strength at room temperature are likely to have reasonable high strength at elevated temperatures. Furthermore, we predicted the strength values of these filtered out candidates at higher temperatures. Without losing the generality of the approach, one could have screened out the alloy candidates at a higher temperature and predicted the strength of these screened out materials in other temperature ranges.

### Multi-objective optimization

Our design problem includes optimization of both yield strength and tensile strength. A pareto front exists where tradeoffs between these parameters can be explored for all non-dominated solutions. However, in this work, we simultaneously optimize for both properties through the construction of two different composite property targets. As per the available experimental niobium alloy datasets, YS and UTS are linearly correlated and have a correlation coefficient value 0.975. We integrated YS and UTS into a composite targeted property and tried with the following formulations:

$$\text{Composite target 1. } y = y_{ys} + y_{uts},$$

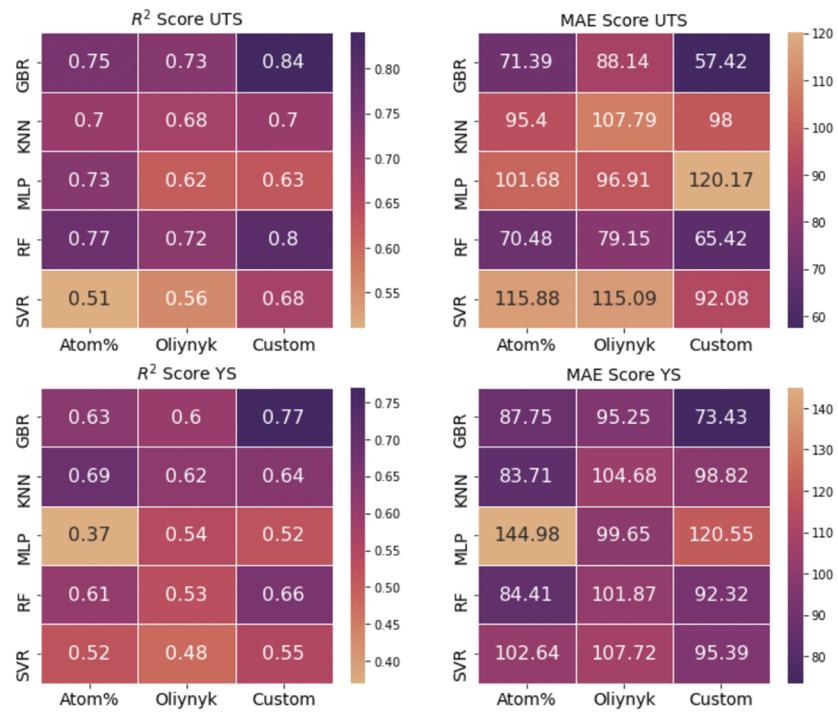
$$\text{Composite target 2. } y = \sqrt{y_{ys}^2 + y_{uts}^2},$$

where  $y_{ys}$  and  $y_{uts}$  represents YS and UTS value of alloy, respectively.

The composite targets are linearly correlated with the YS and UTS with correlation coefficients 0.99+ (Fig. S10, Table TS11, supplementary material). For each of these composite targets, we trained a GBR model and applied bootstrap sampling for estimating the uncertainty as described earlier. Subsequently, promising candidates were screened out using Bayesian optimization.

## RESULTS AND DISCUSSION

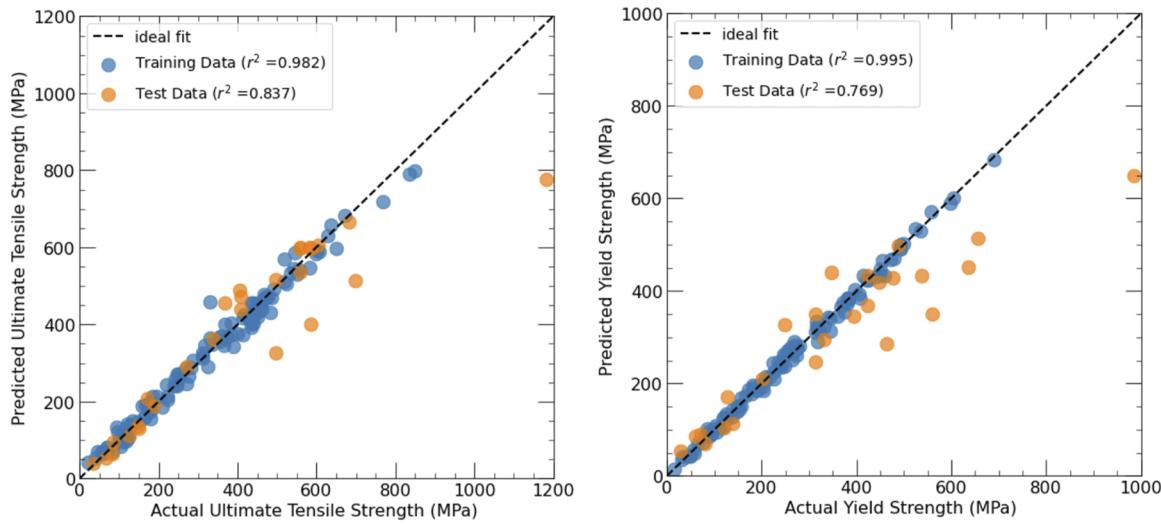
We used GBR, RF, SVR, KNN, and MLP with different featurization schemes to predict UTS and YS values. GBR performed better for YS and UTS in the training and test datasets (Fig. 2 and Tables TS1 and TS2, supplementary material). We have observed that when we used domain knowledge-based custom features, the accuracy of prediction for most of the ML models has been improved for the test dataset. GBR with custom features has the highest  $R^2$



**FIG. 2.** ML models prediction performance for UTS and YS on test dataset using different featurization schemes. Alloys' strength values are predicted at their corresponding available temperature ranging from 24 to 1871 °C.

score of 0.84 and the lowest MAE value of 57.42 MPa for predicting UTS. Similarly, GBR with custom features has the highest R<sup>2</sup> score of 0.77 and the lowest MAE value of 73.43 MPa for predicting YS. With this result, we emphasize that domain knowledge-based

feature selection is an effective compromise between overly simple and overly complex feature sets when a limited dataset is used. Parity plots for strength values of GBR with the custom featurization are shown in Fig. 3. We observed that prediction accuracy



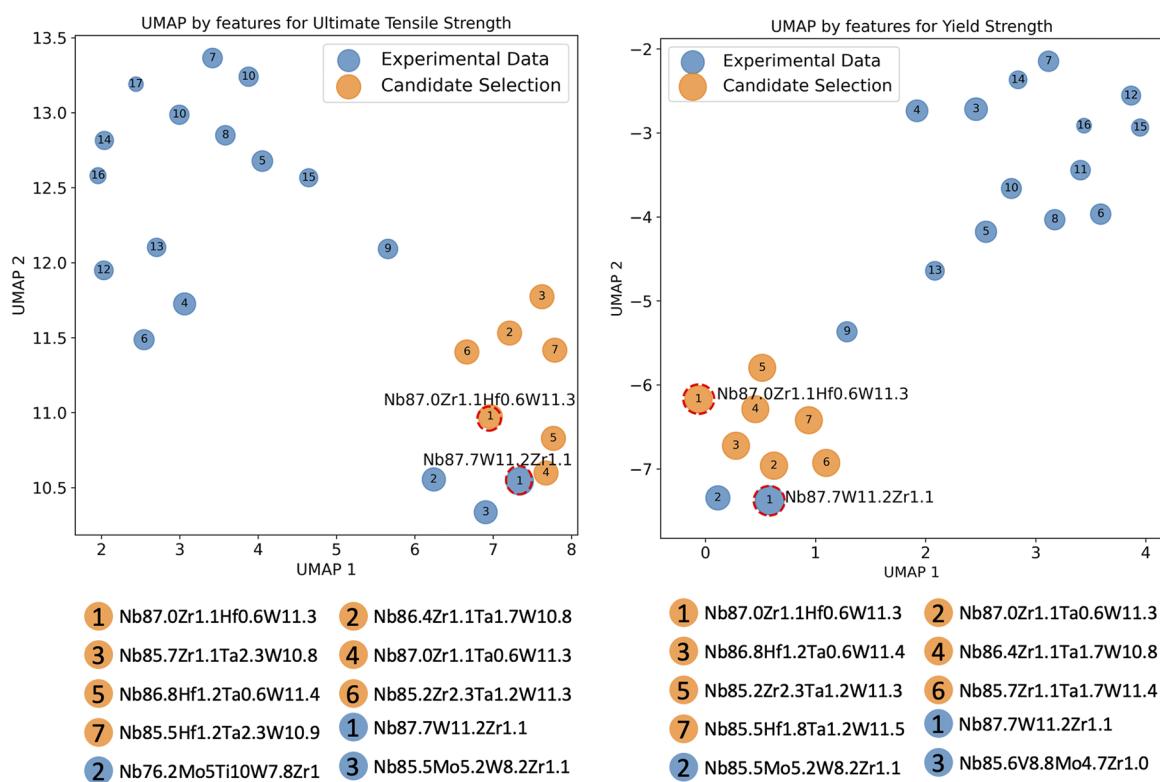
**FIG. 3.** Actual vs predicted ultimate tensile strength (left) and yield strength (right) using GBR model. The predictions of the alloys' strength values are carried out at their corresponding available temperature ranging from 24 to 1871 °C.

at a higher strength value is low because of the data sparsity in the higher strength regime. As GBR outperforms the rest of the models, we continued with GBR as a surrogate model for Bayesian optimization. We used the scikit-learn python library for the implementation of GBR and obtained the feature importance (Tables TS3 and TS4, supplementary material). As expected, the temperature is the key feature for the prediction of strength values as the strength values reduce with increasing temperature. Bulk electron concentration, Pauling electronegativity, and elastic modulus are the most important features in predicting ultimate tensile strength and yield strength. Some features, such as melting point and solid solubility limit, have very little or no significance. As the design strategy is a part of active learning, the screened-out candidates can be experimentally evaluated and added back in subsequent training iterations to get the next set of suggested candidates. If the new alloy candidates will be added to the training dataset, the importance of the features is likely to change. Additionally, some features are highly correlated (Fig. S4, supplementary material), and upon removing these correlated features, we did not observe significant improvement in the model's performance.

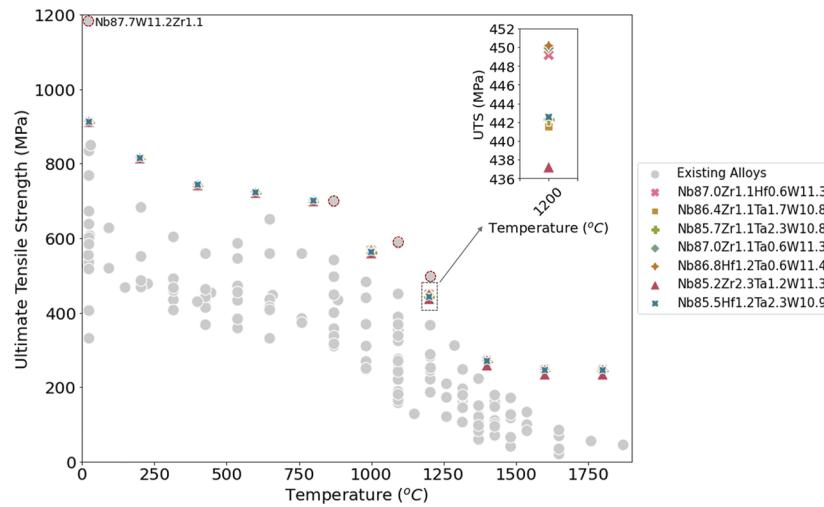
The mean and standard deviation associated with the model predictions were estimated using the bootstrap technique. Figures S5 and S6, supplementary material shows the model performance in the train and test datasets for predicting UTS and YS, respectively.

We noted that the prediction means in the high-strength regions are lower than the actual values and have a high standard deviation. We hypothesize this is due to the lack of sufficient training data in the high-strength region. We estimated the mean and standard deviation of the target properties at room temperature for each alloy composition created in the virtual candidate space and used them in Bayesian optimization to screen out high-performing alloy compositions. We considered the top seven suggested candidates for further analysis (Tables TS5–TS8, supplementary material). The predicted UTS values of the top seven candidates are around 910 MPa with standard deviations ~130 MPa. These suggested quaternary alloy candidates were picked by BO because their mean predictions are near the known highest strength value ( $\text{Nb}_{87.7}\text{W}_{11.2}\text{Zr}_{1.1}$  with 1182 MPa) and have high uncertainty. We may achieve strength values higher than 1182 MPa for these alloys, but it needs experimental validation.

To analyze further, we created the UMAP (Uniform Manifold Approximation and Projection) representations of these seven candidates along with the available known alloys (Fig. 4, Tables TS9 and TS10). The suggested alloy candidates tend to be closer to the cluster of existing high-strength alloys. It seems the algorithm suggested these candidates by exploring the region near to the existing high-strength alloys. However, the model is also capable of exploring the alloy candidates away from the known high-strength alloys.



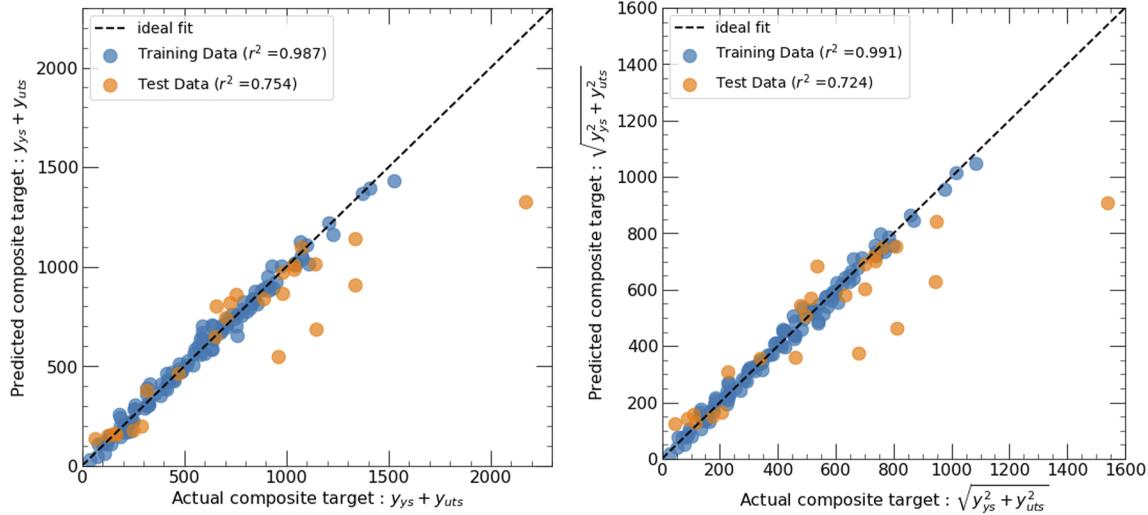
**FIG. 4.** UMAP embeddings of experimental and suggested quaternary alloy compositions for ultimate tensile strength (left) and yield strength (right) where the size of the circle scales with the observed or predicted corresponding value of the mechanical property. The detailed list of the labeled experimental alloys is provided in Tables TS9 and TS10.



**FIG. 5.** Predicted UTS for the screened-out quaternary candidates and observed UTS of the existing Nb alloys at varying temperatures. Inset figure represents predicted UTS values at 1200 °C.

When we used the model for filtering out the quinary alloy candidates, we observed some of the suggested alloys form a cluster away from the existing high-strength alloys (Figs. S7 and S8, supplementary material), these alloys contain rhenium as one of the alloying elements. As we look for materials with higher strength over a broader range of temperatures, we predicted the UTS and YS values of the suggested alloy candidates from room temperature to 1800 °C. Figure 5 shows the predicted UTS for the screened-out quaternary candidates and observed UTS of the existing Nb alloys at varying temperatures. These selected compositions have

superior predicted strength over a range of temperatures (Fig. S9, supplementary material shows results for YS). BO also suggests similar alloy compositions for yield strength; this may be due to the high correlation between YS and UTS (Fig. S10). Thus, we tried combining YS and UTS in 1:1 ratio and formulated the composite target properties. We used GBR with custom features to predict these composite properties, and the model performance is represented in Fig. 6 (Table TS11, supplementary material). The first seven suggested quaternary alloy compositions by BO for the composite properties are presented in Tables TS12 and TS13



**FIG. 6.** Actual vs predicted composite property  $y = y_{ys} + y_{uts}$  (left) and  $y = \sqrt{y_{ys}^2 + y_{uts}^2}$  (right). The predictions of the alloys' composite properties are carried out at their corresponding available temperature ranging from 24 to 1871 °C.

(supplementary material). We observed that the suggested alloys have similar elemental compositions as we observed for YS and UTS individually.

The key alloying elements in the suggested compositions are W, Hf, Zr, Ta, and Re. The maximum solid solubility limit of W, Hf, Zr, and Ta in Nb is 100 at. %, and Re has a solid solubility limit of 45.5 at. %.<sup>19</sup> Thus, the suggested alloy compositions likely form a single-phase solid solution contributing to the alloy strengthening mechanism. However, we have not yet used thermodynamic modeling software to evaluate beyond the binary systems, and it remains a possibility that ternary or more complex alloys could form intermetallic multiphase composites. We used the CALPHAD (CALculation of Phase Diagrams) approach to study equilibrium phase diagrams for these suggested alloy candidates. We constructed pseudo-binary phase diagrams for the suggested alloy candidates  $\text{Nb}_{87}\text{Zr}_{1.1}\text{Hf}_{0.6}\text{W}_{11.3}$  and  $\text{Nb}_{85}\text{Zr}_{1.2}\text{Ta}_{3.5}\text{W}_{9.7}\text{Re}_{0.6}$  using Thermo-Calc software.<sup>22</sup> These phase diagrams are constructed by varying the mole percentage of tungsten while keeping the mole percentage ratio of other elements fixed. The pseudo-binary phase diagrams (Fig. S11, supplementary material) suggest that single-phase BCC solid solution is the only thermodynamically stable phase for a wide temperature range from about 550 to 2500 °C for these alloys. This suggests, the alloy candidates are likely to have high strengths due to the solid solution phase.

## CONCLUSIONS

Nickel superalloys have been used for many decades as turbine materials but are reaching the limits of their utility, and data-driven alloy design is critical for finding entirely new parent alloy systems. We have provided a design strategy for machine-learning guided approach to screen out promising novel niobium alloy candidates as turbine materials from a huge search space. Several well-known machine learning models combined with different material descriptors are tried, and we find that the domain knowledge-based material descriptors are effective for predicting the targeted properties in the case of sparse datasets. We have used Bayesian optimization in filtering out promising novel niobium-based alloy compositions. These compositions have superior predicted strength over a range of temperatures, making them suitable for gas turbines. The identified candidates show the exploitative and explorative capability of the approach in discovering novel materials. We hope this computational approach will complement the experiment and considerably accelerate the development of niobium alloys.

## SUPPLEMENTARY MATERIAL

The supplementary material includes figures of existing niobium alloys temperature vs strength values, working example of different featurization schemes, feature correlation, machine learning models' performance metrics, feature importance, prediction uncertainty, suggested niobium alloy compositions, UMAP representation of quaternary/quinary alloys, predicted YS of the suggested quaternary alloy candidates with varying temperature, and pseudo-binary phase diagram.

## ACKNOWLEDGMENTS

This work was financially supported by the Department of Energy/Advanced Research Projects Agency – Energy (ARPA – E) under Award No. DE-AR0001426.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Trupti Mohanty:** Conceptualization (equal); Data curation (supporting); Investigation (equal); Methodology (equal); Software (lead); Validation (lead); Visualization (lead); Writing – original draft (equal); Writing – review & editing (equal). **K. S. Ravi Chandran:** Data curation (lead); Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal). **Taylor D. Sparks:** Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (lead); Supervision (lead); Writing – original draft (equal); Writing – review & editing (equal).

## DATA AVAILABILITY

The design strategy mentioned in the paper was implemented in Python. Source codes, datasets, and algorithms are available in [https://github.com/truptimohanty/Nballoy\\_BO](https://github.com/truptimohanty/Nballoy_BO).

## REFERENCES

1. J. H. Perepezko, "The hotter the engine, the better," *Science* **326**(5956), 1068–1069 (2009).
2. T. E. Tietz and J. W. Wilson, *Behavior and Properties of Refractory Metals*, 1st ed. (Edward Arnold Publishers, London, 1965), pp. 40–155.
3. J. Wadsworth, T. G. Nieh, and J. J. Stephens, "Recent advances in aerospace refractory metal alloys," *Int. Mater. Rev.* **33**(1), 131–150 (1988).
4. B. A. Wilcox, "Basic strengthening mechanisms in refractory metals," in *Refractory Metal Alloys Metallurgy and Technology* (Springer, Boston, MA, 1968), pp. 1–39.
5. J. Schmidt *et al.*, "Recent advances and applications of machine learning in solid-state materials science," *npj Comput. Mater.* **5**(1), 83 (2019).
6. R. E. A. Goodall and A. A. Lee, "Predicting materials properties without crystal structure: Deep representation learning from stoichiometry," *Nat. Commun.* **11**(1), 6280 (2020).
7. A. Y.-T. Wang *et al.*, "Compositionally restricted attention-based network for materials property predictions," *npj Comput. Mater.* **7**(1), 77 (2021).
8. A. Mansouri Tehrani *et al.*, "Machine learning directed search for ultraincompressible, superhard materials," *J. Am. Chem. Soc.* **140**(31), 9844–9853 (2018).
9. X. Liu *et al.*, "Material machine learning for alloys: Applications, challenges and perspectives," *J. Alloys Compd.* **921**, 165984 (2022).
10. C. Wang *et al.*, "A property-oriented design strategy for high performance copper alloys via machine learning," *npj Comput. Mater.* **5**(1), 87 (2019).
11. A. Nazarabari and D. Canadinc, "Prediction of the NiTi shape memory alloy composition with the best corrosion resistance for dental applications utilizing artificial intelligence," *Mater. Chem. Phys.* **258**, 123974 (2021).
12. C. Wen *et al.*, "Machine learning assisted design of high entropy alloys with desired property," *Acta Mater.* **170**, 109–117 (2019).

- <sup>13</sup>K. S. Ravi Chandran, *Compilation of Niobium Alloy Mechanical Properties* (Department of MSE, University of Utah, Salt Lake City, UT, 2021).
- <sup>14</sup>D. Jha *et al.*, “ElemNet: Deep learning the chemistry of materials from only elemental composition,” *Sci. Rep.* **8**(1), 17593 (2018).
- <sup>15</sup>S. Kirklin *et al.*, “The open quantum materials database (OQMD): Assessing the accuracy of DFT formation energies,” *npj Comput. Mater.* **1**(1), 15010 (2015).
- <sup>16</sup>R. J. Murdock *et al.*, “Is domain knowledge necessary for machine learning materials properties?,” *Integr. Mater. Manuf. Innovation* **9**(3), 221–227 (2020).
- <sup>17</sup>A. Y.-T. Wang *et al.*, “Machine learning for materials scientists: An introductory guide toward best practices,” *Chem. Mater.* **32**(12), 4954–4965 (2020).
- <sup>18</sup>S.-A. Cho, “Bulk electron concentration and average valence bond strength of elements,” *Trans. Jpn. Inst. Met.* **22**(9), 643–652 (1981).
- <sup>19</sup>R. Goodall, “Data of the maximum solid solubility limits of binary systems of elements,” *Data Brief* **26**, 104515 (2019).
- <sup>20</sup>L. Mentel (2014). “mendeleev—A Python resource for properties of chemical elements, ions and isotopes,” GitHub, version 0.9.0. <https://github.com/lmentel/mendeleev>.
- <sup>21</sup>F. Nogueira (2014). “Bayesian optimization: Opensource constrained global optimization tool for Python,” GitHub. <https://github.com/fmfn/BayesianOptimization>.
- <sup>22</sup>SSOL6: SGTE Alloy Solutions Database v6.0, Thermo-Calc software 2021b; accessed 3 August 2022.