Mhadeshwar *et al. J Mater Inf* 2024;4:12
DOI: 10.20517/jmi.2024.17

Journal of Materials Informatics

Research Article

Open Access

Check for updates

Borides as promising M_2AX phase materials with high elastic modulus using machine learning and optimization

Ashwin Mhadeshwar, Trupti Mohanty, Taylor D. Sparks*

Department of Materials Science & Engineering, University of Utah, Salt Lake City, UT 84112, USA.

***Correspondence to:** Prof. Taylor D. Sparks, Department of Materials Science & Engineering, 122 S. Central Campus Drive, University of Utah, Salt Lake City, UT-84112, USA. E-mail: sparks@eng.utah.edu

How to cite this article: Mhadeshwar A, Mohanty T, Sparks TD. Borides as promising M_2AX phase materials with high elastic modulus using machine learning and optimization. *J Mater Inf* 2024;4:12. https://dx.doi.org/10.20517/jmi.2024.17

Received: 20 Jun 2024 **First Decision:** 24 Jul 2024 **Revised:** 21 Aug 2024 **Accepted:** 23 Aug 2024 **Published:** 27 Aug 2024

Academic Editor: Xingjun Liu **Copy Editor:** Dong-Li Li **Production Editor:** Dong-Li Li

Abstract

There is growing interest in novel MAX phase materials for various applications ranging from aircraft/spacecraft and defense to energy and electronics due to their unique combination of metallic and ceramic properties. Traditional materials discovery has mostly relied on human intuition coupled with rigorous experiments; however, this approach has been time-consuming and inefficient. Over the last few decades, advances in fundamental and data-driven approaches such as first-principles modeling, materials informatics, machine learning and optimization, coupled with an exponential rise in computational power, have enabled faster and more efficient materials discovery. Here, we present an exploration of high elastic modulus novel boride-based M_2AX phase materials using a combination of the aforementioned methods. Specifically, an ensemble of gradient boosted machine learning models was developed to predict the elastic modulus from informatics-based structural features by leveraging a dataset of Density Functional Theory (DFT)-predicted elastic moduli for 223 M_2AX phase materials (carbides and nitrides). Using Bayesian optimization, inverse modeling was carried out to maximize the model-predicted elastic modulus by identifying the optimal features. Finally, model predictions for 1,035 candidate M_2AX materials were generated to compare their features with the optimal features to identify potential novel promising materials. We found that Ta_2PB, Nb_2PB, and V_2PB have similar high elastic moduli (371.7, 351.5, and 347.4 GPa) to their carbide counterparts (364.7, 357.7, and 373.5 GPa), and our results support the possibility that borides can be a viable tertiary element for M_2AX phases.

Keywords: M_2AX phase, materials informatics, machine learning, ensemble, Bayesian optimization, borides

INTRODUCTION

MAX phases are a group of materials that have drawn increasing interest in recent years due to their unique combination of metallic and ceramic properties[1]. These materials tend to follow a composition of $M_{n+1}AX_n$ and are typically carbides and nitrides with a layered, ternary composition[2,3]. MAX phases exhibit ceramic properties through their high strength and resistance to oxidation up to 1,400 °C, and high elastic stiffness[4,5]. Additionally, they display metallic properties of thermal and electrical conductivity, machinability, and high damage tolerance[1,4,6]. Together, these properties make MAX phases highly promising nanolaminates for aerospace, defense, and energy applications. Specifically, MAX phases can act as structural materials or coatings for high-temperature applications such as gas turbine engines or gas burner nozzles due to their ability to withstand high temperatures for long periods of time[4,7]. Due to their superior oxidation and corrosion resistance and radiation tolerance, MAX phases have also shown potential as accident-tolerant fuel cladding in nuclear reactors[8]. Chirica *et al.* reported that when used as catalysts, MAX phases can activate C–H or C–O bonds in various chemical reactions such as methanol formation, dry reforming, CO oxidation or water-gas shift reactions[9]. Given the diverse applications of MAX phases, they are indeed worthy of further exploration for new materials discovery.

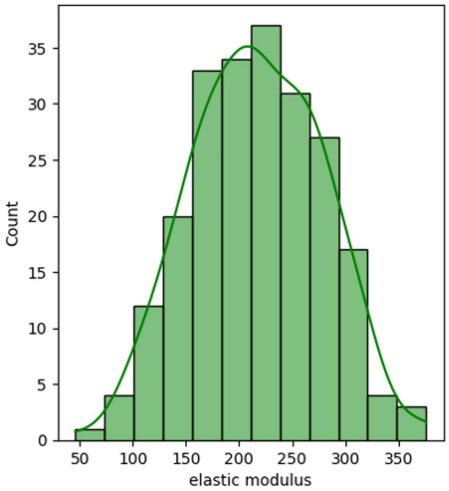
In the past, materials science has relied on making small structural or compositional adjustments to known materials in order to discover novel alternatives. While this method has produced results, it is costly, time-consuming, and inefficient. In recent decades, the process of materials discovery has been accelerated due to the increase in computational processing power, formation and availability of materials databases, and development of new modeling/data science methodologies[10-20]. In terms of methodology development, Density Functional Theory (DFT) simulations have long-established the electronic structure of materials and their properties. Use of machine learning (ML) and optimization approaches to the design of novel materials has been a groundbreaking shift in the field[21]. ML has already been applied for the development of novel materials in numerous domains and has produced exceptional results. For example, Li *et al.* utilized it to find an aluminum-lithium alloy with a 12.6% higher specific modulus and a similar specific strength compared to the typically used 2195-T8 alloy[22]. Similarly, Mohanty *et al.* explored niobium-based alloys through ML and Bayesian optimization (BO) and found two new niobium alloys that maintain high strength from 550 to 2,500 °C based on calculation of phase diagrams[23].

In the present study, we utilized ML and optimization approaches to search for new M_2AX phase materials with high elastic moduli which could be used for future materials development. Specifically, we utilized a Matminer dataset generated by Cover *et al.* which provided data on 223 M_2AX phase materials including elastic, shear, and bulk moduli, elastic constants, distances between their atoms, lattice parameters, and chemical formulas[24]. We applied Matminer featurizers to the compositional formulas to generate various structural features. We developed an ensemble of gradient boosted ML models to predict the elastic modulus from those informatics-based structural features. We also carried out inverse modeling using BO to maximize the model-predicted elastic modulus by identifying the optimal features. Finally, we generated model predictions for 1,035 various M_2AX materials to compare their features with the optimal features to identify potential novel materials. Our work indicates that Ta_2PB, Nb_2PB, and V_2PB have similar high elastic moduli to their carbide counterparts, which supports the possibility that borides can be a viable tertiary element for M_2AX phases.

MATERIALS AND METHODS

Preliminary data analysis and feature creation

The M_2AX dataset used in this work included elastic modulus data from 223 DFT simulations conducted by Cover *et al.* A histogram and the statistics of the elastic modulus data are shown in Figure 1[24]. The data



	Elastic Modulus (GPa)
Mean	217
Standard Deviation	61
Minimum	46
25th Percentile	173
Median	217
75th Percentile	263
Maximum	376

Figure 1. Histogram and summary statistics for elastic modulus data from the DFT dataset for the 223 M_2AX materials[24]. DFT: Density Functional Theory.

ranges from 46 to 376 GPa, with a mean of 217 GPa, a standard deviation of 61 GPa, a median of 217 GPa, and an interquartile range (IQR) of 90 GPa. The wide range and fairly high standard deviation provide an excellent opportunity to capture the diversity in this elastic modulus dataset through ML modeling. The Shapiro-Wilk normality test indicated that the data closely resembles the Gaussian distribution (P = 0.667). Other information in the dataset includes structural parameters (a: Lattice parameter a in angstroms, c: lattice parameter c in angstroms, 5 elastic constants of the M_2AX material specific to hexagonal materials, d_ma: distance from M atom to A atom, d_mx: distance from M atom to X atom) and the formula.

Using the StrToComposition Matminer featurizer[10], formulas from the original dataset were converted to a usable format; e.g., Sc_2AlC is converted to (Sc, Al, C). The developed ML model is specifically applicable to predicting the elastic modulus of the 211 MAX phase and does not account for other compositions/phase effects. Next, the ElementProperty Matminer featurizer[10] was used to generate 132 structural features for each compound. Examples of these features include range GSmagmom, minimum SpaceGroupNumber, and mean NUnfilled. The original parameters or features provided in the DFT dataset were not used as the goal was for the model to be able to predict the elastic modulus of novel materials beyond this dataset.

Feature selection and ensemble of models

Data was split into 80% training and 20% validation. Initially, an XGBoost (XGB) regressor model was developed with 5-fold cross validation using all 132 features. Hyperparameter optimization was conducted using the number of estimators, learning rate, maximum depth, and minimum loss reduction. The maximum depth was restricted to 8 to avoid overfitting. Performance of the XGB models is influenced by the random state used to split the data into training and validation; hence, four different random states (42, 0, 14, and 28) were used to develop four separate XGB models. Permutation importance for each model was calculated and is shown in Figure 2 for the top 15 features. Starting from the highest importance values, models were developed with an increasing number of features until no significant improvement in validation root mean square error (RMSE) and R^2 was observed [Figure 3]. While selecting the features for each model, any additional feature with high correlation (threshold = 0.8) with previously included features was not considered. Typically, performance of these XGB models reaches a plateau after including 6-8 features. Finally, a list of eight features was selected based on their common appearance across the four models, and the models were re-developed using those eight features at the four different random states. The final list of selected features includes: (1) maximum GSvolume_pa, (2) mean MendeleevNumber, (3)

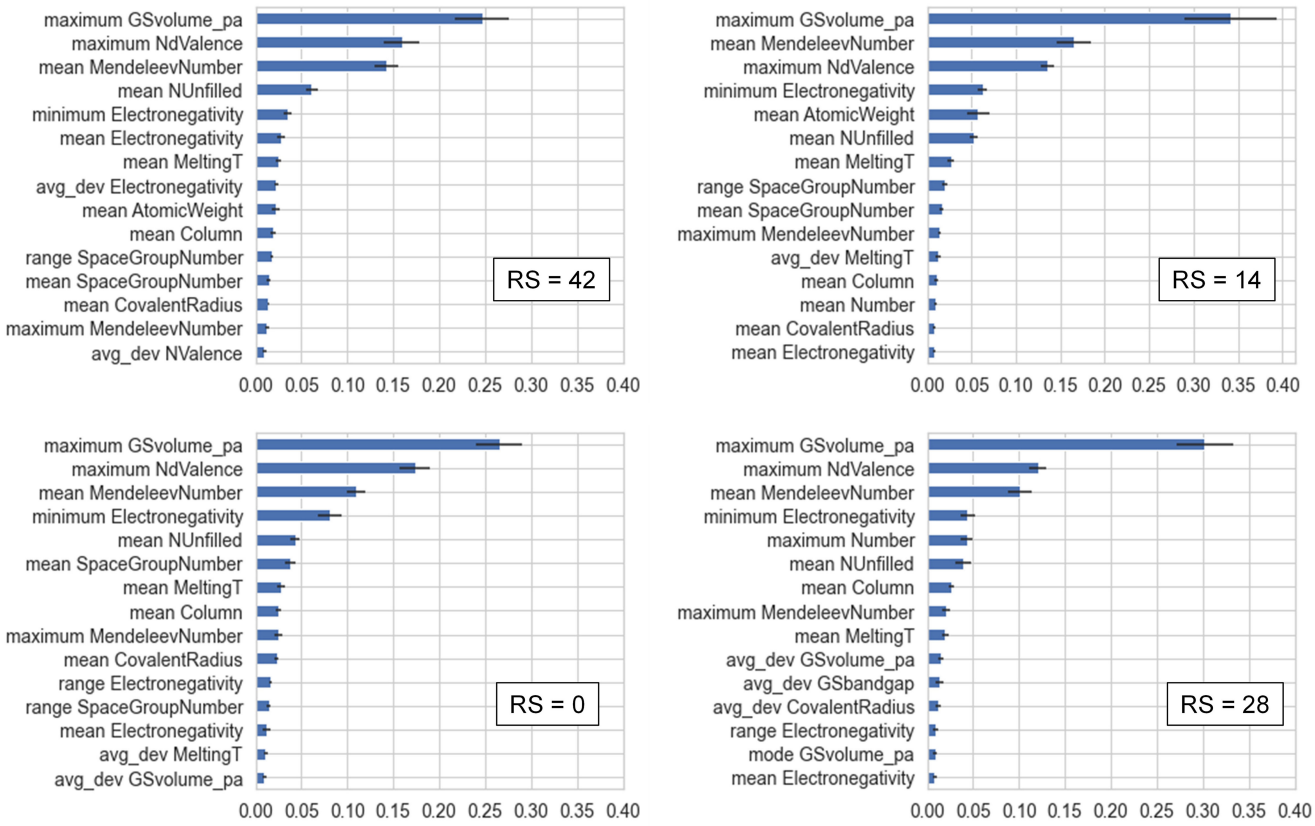


Figure 2. Permutation importance plots for the top 15 features for XGB models developed using all 132 features with four random states (RS = 42, 0, 14, 28) to predict the elastic modulus for M_2AX materials. These importances were used to explore stepwise addition of features for each model. XGB: XGBoost; RS: random states.

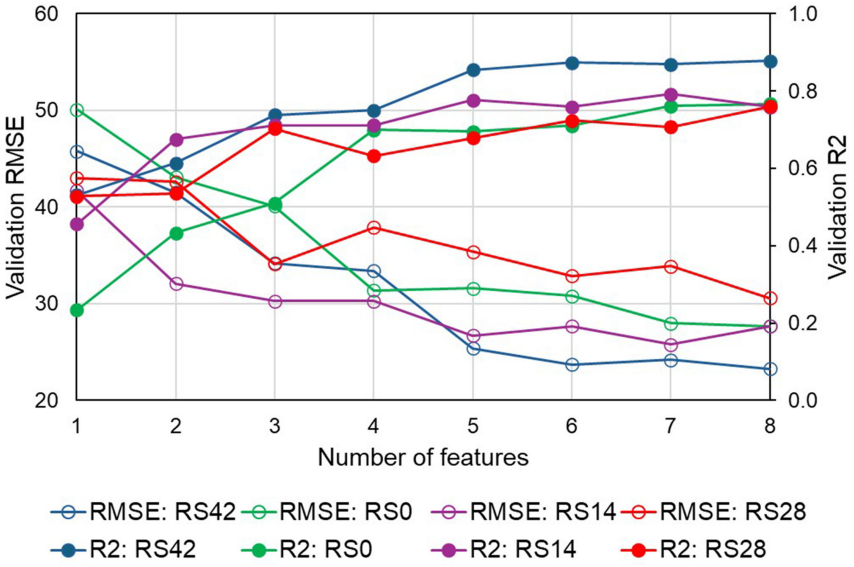
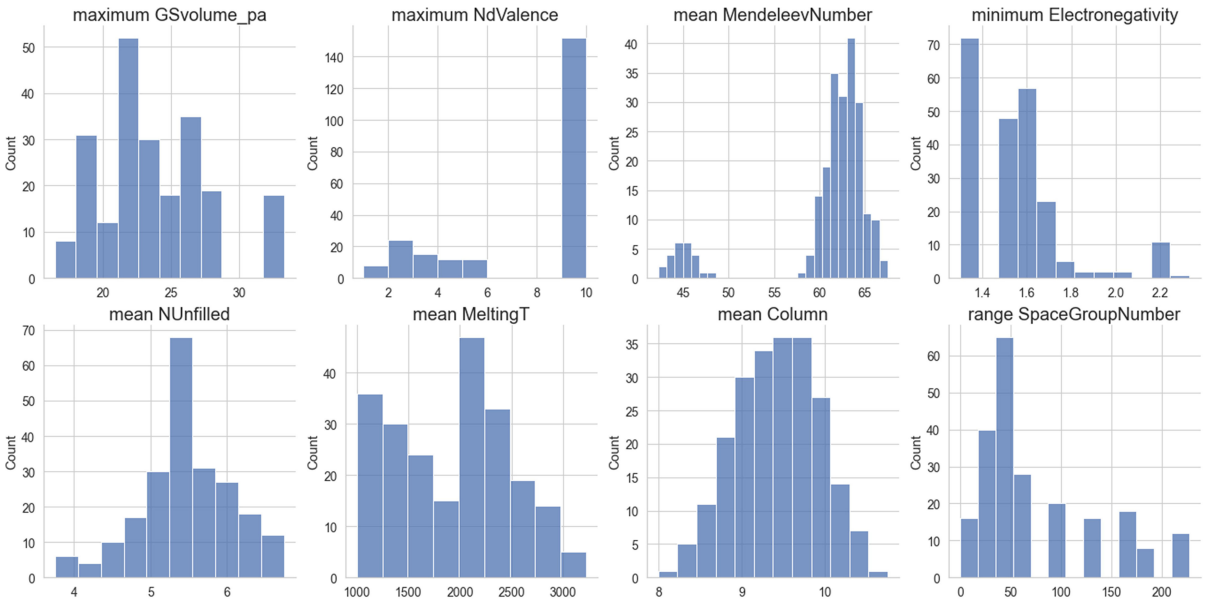


Figure 3. Effect of stepwise feature addition on XGB model performance in terms of validation RMSE and R^2. Four random states (RS = 42, 0, 14, and 28) are shown. Note that the final four models [Table 1] were re-developed using a list of eight features selected based on their common appearance across the four models shown here. XGB: XGBoost; RMSE: root mean square error; RS: random states.



	Maximum GSvolume_pa	Maximum NdValence	Mean MendeleevNumber	Minimum Electronegativity	Mean NUnfilled	Mean MeltingT	Mean Column	Range SpaceGroupNumber
Mean	23.84	7.753	60.85	1.551	5.441	1931.3	9.425	72.31
Standard Deviation	4.038	3.373	5.797	0.225	0.614	569.0	0.548	61.95
Minimum	16.48	1	42.25	1.30	3.75	998.5	8	0
25th Percentile	22.18	4	61	1.36	5	1388.8	9	34
Median	23.01	10	62.5	1.54	5.5	2011.8	9.5	35
75th Percentile	26.91	10	63.75	1.63	6	2349.4	9.75	90
Maximum	33.29	10	67.5	2.33	6.75	3225	10.75	227

Figure 4. Histogram and summary statistics for the final eight selected features used in model development.

maximum NdValence, (4) minimum Electronegativity, (5) mean NUnfilled, (6) mean MeltingT, (7) mean Column, and (8) range SpaceGroupNumber. The distribution and statistics of these eight features are shown in Figure 4 and their cross-correlations are shown in Figure 5.

RESULTS AND DISCUSSION

Parity plots for the four models are shown in Figure 6 for the validation datasets. Parity plots for these models and their ensemble covering the entire dataset (training and validation) are provided in the Supplementary Figure 1. This unique approach of selecting features from a collection (ensemble) of models is expected to make the overall model robust, avoid overfitting, and capture variations in the overall dataset. A further advantage of developing four ML models (instead of a single model that is typically developed) is that the ensemble of these four models can now be used to improve overall prediction accuracy, while ensuring that the overall model has a wide grasp of the majority of the dataset. The ensemble approach averages predictions from these models to produce a final prediction for any given material, which is expected to improve the robustness of the predictions, dampen the effect of outlier predictions from any single model, and make the predictions less sensitive to any one specific way of data splitting. Details of the four XGB models (data splitting, hyperparameters, and metrics for validation) are presented in Table 1. We acknowledge that the development of multiple (ensemble) ML models, although desired, can be time-consuming for large datasets.

Table 1. Hyperparameters and metrics for the ensemble of XGB models developed at multiple random states to predict elastic modulus for M_2AX materials

Category	Description	Model 1	Model 2	Model 3	Model 4
Data splitting	Training: validation	80:20	80:20	80:20	80:20
	Random state	42	0	14	28
Hyperparameters	Number of estimators	75	100	100	100
	Learning rate	0.1	0.1	0.1	0.1
	Maximum depth	6	6	6	6
	Minimum split loss	5	1	5	3
Metrics for validation	R^2	0.892	0.790	0.795	0.721
	MAE	16.9	19.6	20.2	25.3
	MSE	484	690	653	1,090
	RMSE	22.0	26.3	25.5	33.1

XGB: XGBoost; MAE: mean absolute error; MSE: mean squared error; RMSE: root mean squared error.

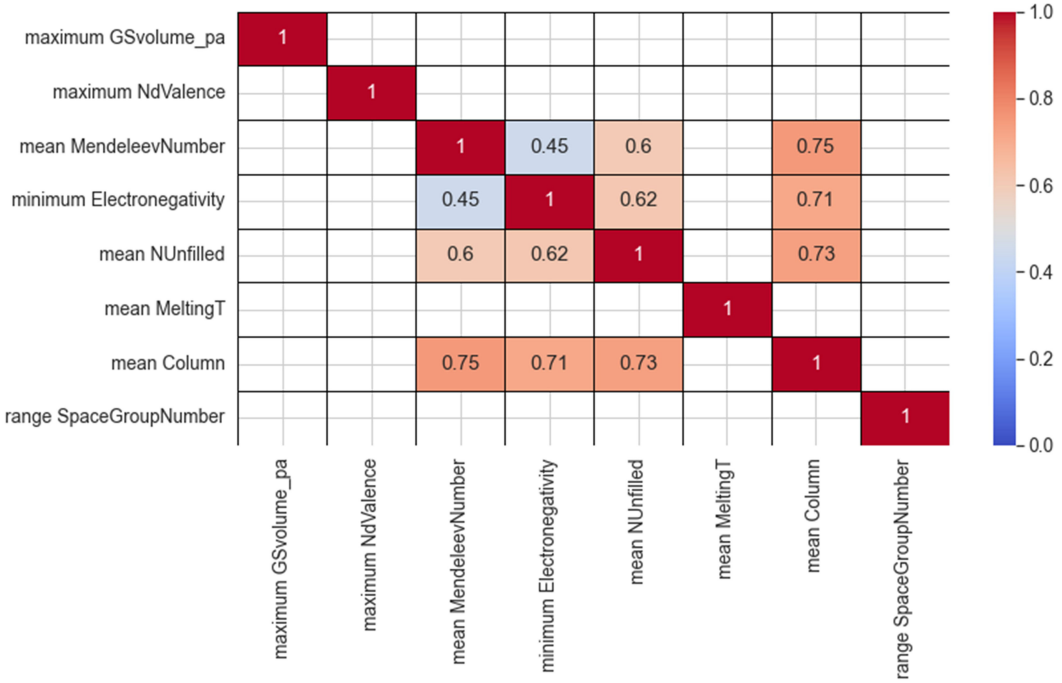


Figure 5. Cross-correlations among the final eight selected features used in the development of the XGB models. XGB: XGBoost.

Finally, Figure 7 shows the partial dependence plots generated using the ensemble of XGB models. Here, each of the eight features was varied from its minimum to maximum value in 30 steps while keeping the seven remaining features at their mean values. Although there might be restrictions on whether the features can be varied independently, this type of analysis can provide helpful directions toward identifying better materials. For example, low maximum GSvolume_pa, low maximum NdValence, moderate mean MendeleevNumber, low minimum Electronegativity, high mean NUnfilled, high mean MeltingT, high mean Column, and high range SpaceGroupNumber might result in M_2AX materials with high elastic modulus.

Bayesian optimization to identify best feature values

To explore promising M_2AX materials with high elastic moduli, we implemented BO using the gp_minimize function in scikit-opt[25]. Here, the objective was to identify feature values corresponding to the maximum predicted elastic modulus. An ensemble of the four XGB models was used in the objective

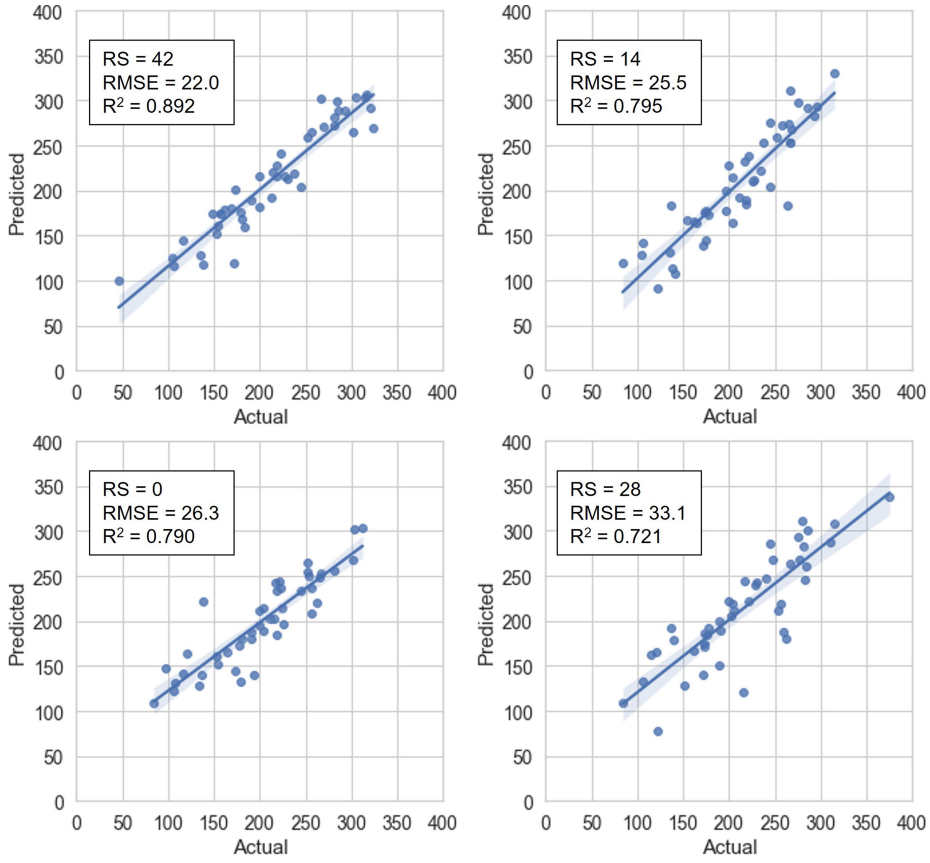


Figure 6. Parity plots for validation datasets using the ensemble of XGB models developed at RS of 42, 0, 14, and 28. XGB: XGBoost; RS: random states; RMSE: root mean square error.

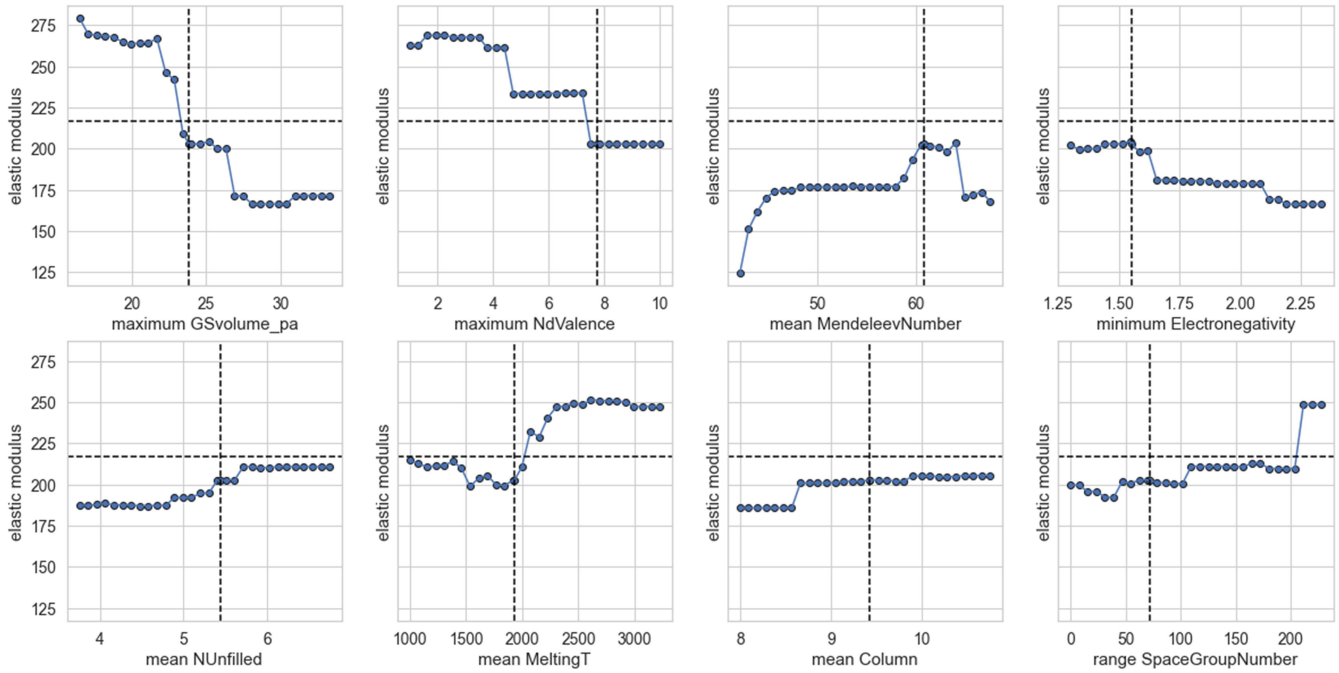
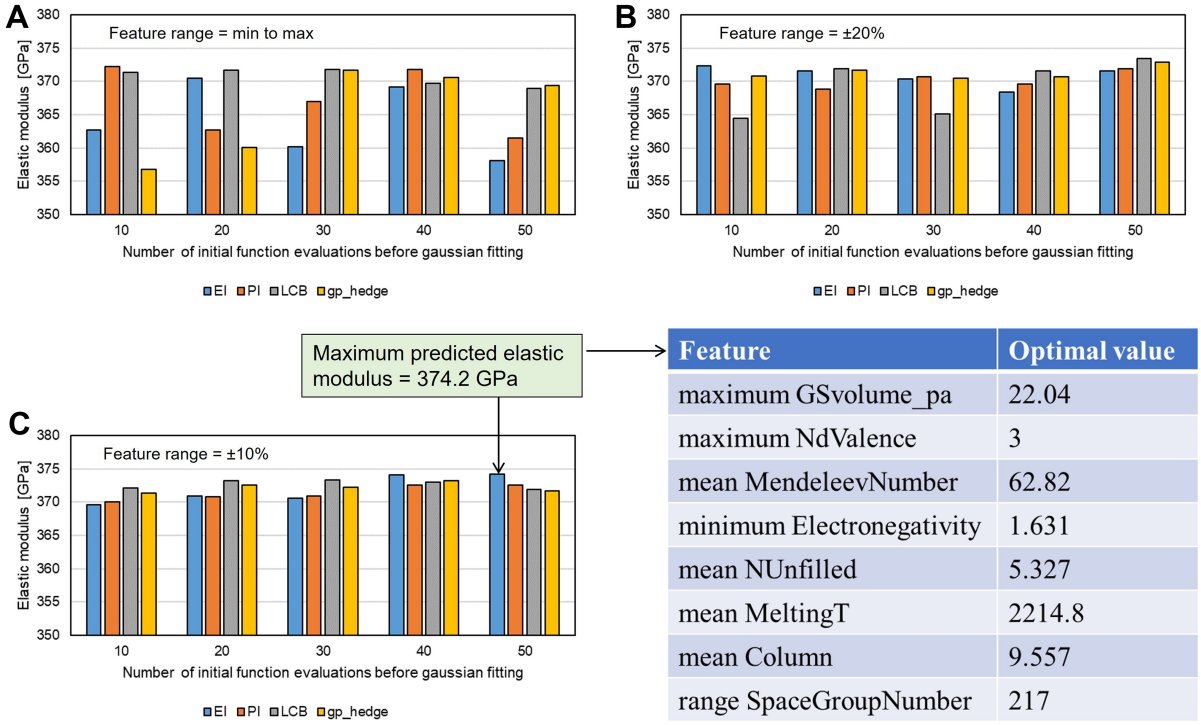


Figure 7. Partial dependence plots showing the effect of each feature on predicted elastic modulus using the ensemble of XGB models. Dashed lines indicate the mean values for the dataset. XGB: XGBoost.



Feature	Optimal value
maximum GSvolume_pa	22.04
maximum NdValence	3
mean MendeleevNumber	62.82
minimum Electronegativity	1.631
mean NUnfilled	5.327
mean MeltingT	2214.8
mean Column	9.557
range SpaceGroupNumber	217

Figure 8. Comparison of acquisition functions and number of initial function evaluations in Bayesian optimization with an ensemble of XGB models, with various ranges (or bounds) for the features (A) min to max, (B) ± 20% from V_2PC, and (C) ± 10% from V_2PC. Features corresponding to the maximum predicted elastic modulus are shown in the table. XGB: XGBoost; EI: expected improvement; PI: probability improvement; LCB: lower confidence bound.

function to get an average elastic modulus. As the gp_minimize function minimizes the objective function, the model prediction was subtracted from 0 inside the objective function. Note that at every iteration of BO, the values that the optimizer generated for maximum NdValence and range SpaceGroupNumber features were typecasted as integers since they are supposed to be represented as integers in the data. As shown in Figure 8, we explored various acquisition functions [expected improvement (EI), probability improvement (PI), lower confidence bound (LCB), gp_hedge] and a range of initial points prior to the Gaussian fitting. For each acquisition function, five runs were completed with the number of initial points ranging from 10 to 50. The number of function evaluations after the initial calls was kept constant at 150; therefore, the total number of function evaluations is equal to 150 + the number of initial points. There was no clear pattern in terms of the acquisition function used, but typically the optimizer has a better chance of finding an optimum with more initial points. With the full range for each feature from its minimum to maximum value, the optimizer resulted in the best prediction of 371.7 GPa, which was slightly lower than the highest elastic modulus predicted from the dataset (373.5 GPa for V_2PC). Therefore, narrower feature ranges around V_2PC were also explored, which identified improved maximum values for elastic modulus (± 20% resulted in 373.4 GPa, whereas ± 10% resulted in 374.2 GPa). The corresponding best feature values are also shown in Figure 8.

Model-driven exploration of novel materials

To identify M_2AX materials with high elastic moduli, we generated 1035 combinations of M, A, and X elements. Elements were selected for each category based on their proximity to commonly used elements in M_2AX phases. We chose 15 early transition metals (M = V, W, Ta, Zr, Hf, Nb, Mo, Cr, Ti, Sc, Y, Mn, Re, Tc, La), 23 group A elements (A = Al, P, Te, Si, Ga, Ge, As, Cd, Sn, In, Tl, Pb, S, Sb, Bi, Zn, Se, Fe, Co, Ni, Cu,

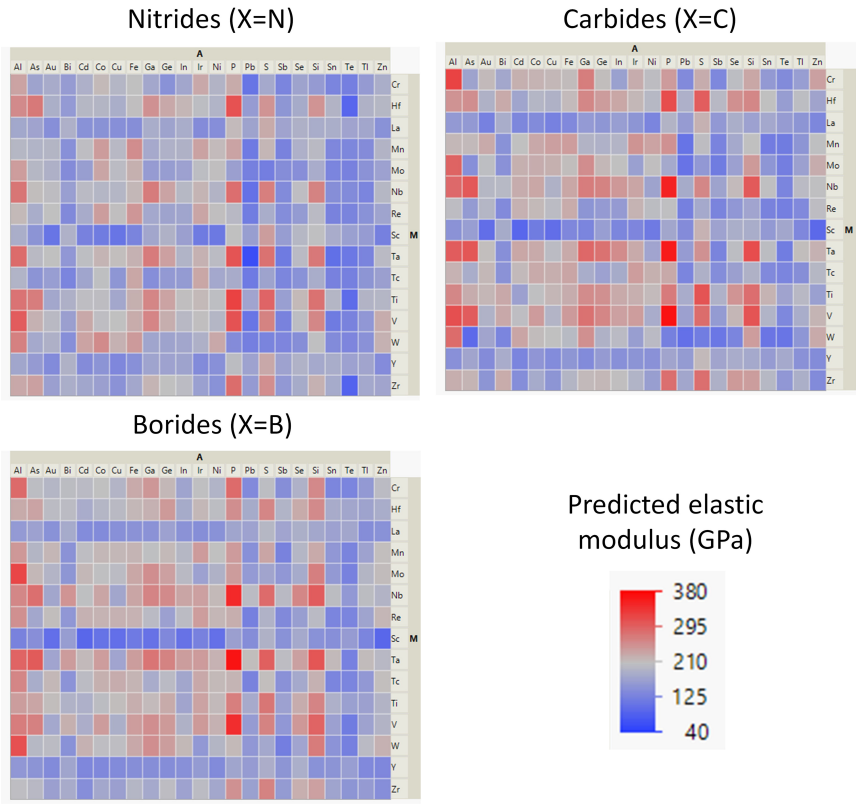


Figure 9. Heatmap representation of the predicted elastic modulus values for screened 1,035 M_2AX materials using the ensemble of XGB models. XGB: XGBoost.

Au, Ir), and three reactive nonmetals / metalloids (X = N, C, B). The same featurization process was used to create the 132 features for the 1,035 materials; however, only the eight selected features in the ML models are used in the analysis presented below.

As a first step, the best feature values from BO were compared against the feature values for all 1,035 materials. A least square difference between the BO feature values and materials features values was calculated for each feature, and then summed over all features to indicate a normalized difference estimate for each material. A small difference is supposed to indicate a potentially promising material with a high elastic modulus. Top ten materials with smallest differences include (V_2PC - 0.005, Ta_2PB - 0.012, V_2PB - 0.047, Ta_2PC - 0.054, Ta_2PN - 0.063, V_2SC - 0.108, Ta_2SB - 0.111, Nb_2PB - 0.122, Nb_2PC - 0.123, V_2SB - 0.140, and Ti_2PC - 0.145). This shows that features for V_2PC are closest to the BO result, but other materials are also promising. Just for relative comparison, the difference was as high as 7.571. Furthermore, five out of the top ten materials are borides, which suggests that some of these borides may potentially have high elastic moduli.

The ensemble of XGB models were then used to generate predictions for the 1,035 materials. A heatmap of the predicted values is shown in Figure 9, where the information is separated into three different types of materials - Nitrides, Carbides, and Borides. Top ten materials with high elastic moduli include V_2PC - 373.5 GPa, Ta_2PB - 371.7 GPa, Ta_2PC - 364.7 GPa, Nb_2PC - 357.7 GPa, Nb_2PB - 351.5 GPa, V_2PB - 347.4 GPa, Ti_2PN - 328.8 GPa, Cr_2AlC - 326.4 GPa, Mo_2AlB - 323.3 GPa, and Hf_2PC - 318.1 GPa, where again four out of the top ten materials are borides. Table 2 shows the comparison of a handful of such

Table 2. Comparison of materials with high elastic modulus based on theory (DFT[24,26]), BO prediction, and machine learning models (ensemble of XGB models)

Features/Target	Best compound from DFT: V_2PC	2nd best compound from DFT: Ta_2PC	3rd best compound from DFT: Nb_2PC	Optimal values from Bayesian optimization	Best new compound from ML exploration: Ta_2PB	2nd best new compound from ML exploration: Nb_2PB	3rd best new compound from ML exploration: V_2PB
Maximum GsVolume_pa	22.570	22.570	22.570	22.04	22.570	22.570	22.570
Maximum NdValence	3	3	4	3	3	4	3
Mean Mendeleev number	63	64	63.5	62.82	62.75	62.25	61.75
Minimum electronegativity	1.63	1.5	1.6	1.631	1.5	1.6	1.63
Mean nunfilled	5.25	5.25	5.25	5.327	5.5	5.5	5.5
Mean MeltingT	2,126.6	2,680.1	2,410.1	2,214.8	2,311.3	2,041.3	1,757.8
Mean Column	9.75	9.75	9.75	9.557	9.5	9.5	9.5
Range space group number	227	227	227	217	227	227	227
Elastic modulus (GPa)	376 (DFT[24]) 373.5 (ML)	375 (DFT[24]) 364.7 (ML)	358 (DFT[24]) 355 (DFT[26]) 357.7 (ML)	374.2 (BO)	371.7 (ML)	351.5 (ML) 355 (DFT[26])	347.4 (ML)

DFT: Density Functional Theory; BO: Bayesian optimization; XGB: XGBoost; ML: machine learning.

promising materials in terms of their features and elastic moduli - either estimated using DFT or predicted using the ensemble of ML models or from BO. Overall, this work indicates that Ta_2PB might be an interesting M_2AX material to explore, and given the uncertainty in ML predictions, Nb_2PB and V_2PB might be promising as well. Recent DFT work by Ali *et al.* also shows that both Nb_2PB and Nb_2PC have a high elastic modulus of 355 GPa, which is consistent with our ML predictions shown in Table 2[26]. It is also noted that although Ti_2PB and Ti_2PC as well as Hf_2SB and Hf_2SC were not found to be the top candidates for elastic modulus, our ML predictions (Ti_2PB = 271.3 GPa, Ti_2PC = 274.2 GPa, Hf_2SB = 264.5 GPa, Hf_2SC = 302.2 GPa) are consistent with the DFT predictions of Ali *et al.* (Ti_2PB = 267 GPa, Ti_2PC = 282 GPa, Hf_2SB = 268 GPa, Hf_2SC = 344 GPa)[26]. Similarly, our ML predictions (Nb_2SB = 289.5 GPa, Zr_2SB = 267.8 GPa, Hf_2SB = 264.5 GPa) are consistent with the DFT predictions (Nb_2SB = 300 GPa, Zr_2SB = 238 GPa, Hf_2SB = 250 GPa) of Zhang *et al.*[27]. Researchers have recently reported synthesis of MAX phase borides[28,29] where the synthesized Nb_2SB, Zr_2SB, and Hf_2SB materials showed the typical $P6_3/mmc$ MAX phase crystal structure, same as that of carbides (e.g., Cr_2AlC).

Finally, we also evaluated and confirmed the stability of these borides based on their predicted energy above convex hull[15,16] using CrabNet. CrabNet is a compositionally restricted attention-based network that predicts material properties, including the energy above the convex hull, using only compositional information[16]. Since we do not have the structural information for the suggested MAX phase borides (Ta_2PB, Nb_2PB, and V_2PB) with high predicted elastic modulus values, we relied on CrabNet for these predictions. Our findings indicate that these borides have relatively low energies above the convex hull (Ta_2PB = 0.087 meV/atom, Nb_2PB = 0.024 meV/atom, and V_2PB = 0.168 meV/atom), suggesting they are among the stable boride candidates. However, a more rigorous and accurate evaluation of these materials' stability could be achieved by theoretically computing the formation enthalpy[3,30-32]. Dahlqvist *et al.* discuss the stability of these phases based on DFT calculations, showing that Ta_2PB and Nb_2PB have formation enthalpies (ΔHcp) of less than 50 meV/atom, indicating they are relatively stable[3]. These findings suggest that further investigation is needed, particularly for Ta_2PB and Nb_2PB.

CONCLUSIONS

In this work, we present an approach for identifying novel boride-based M_2AX phase materials with high elastic modulus using a combination of materials informatics, machine learning and optimization. Starting from a dataset of DFT-predicted elastic moduli for 223 M_2AX phase materials (carbides and nitrides)[24], structural features were generated from composition, and an ensemble of gradient boosted machine learning models were developed to predict the elastic modulus. Inverse modeling was carried out using Bayesian optimization to maximize the model-predicted elastic modulus by identifying the optimal features. Finally, model predictions for 1,035 M_2 AX materials were generated to compare their features with the optimal features to identify potential novel materials. Based on our model predictions, Ta_2PB, Nb_2PB, and V_2PB have similar high elastic moduli to their carbide counterparts, and these materials demonstrate stability based on their predicted energy above convex hull. These findings on borides as a viable tertiary element for M_2AX phases should be further experimentally explored by materials researchers for various applications. The approach presented here, although specific to 211 MAX phase materials, is applicable and could be extended to other MAX phase materials as additional datasets become available.

DECLARATIONS

Authors' contributions

Conceptualization, data curation, investigation, methodology, validation, writing - original draft and revisions: Mhadeshwar A
Methodology, writing - review: Mohanty T
Supervision, methodology, writing - review: Sparks TD

Availability of data and materials

Python source codes and datasets are available at https://github.com/ashwinbmhadeshwar/M2AX_ElasticModulus_MLOPT.

Financial support and sponsorship

The authors gratefully acknowledge funding from the Army Research Office Materials Design program under contract number #W911NF-23-1-0333.

Conflicts of interest

Sparks TD is a Junior Editorial Board Member of *Journal of Materials Informatics*, while the other authors have declared that they have no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. Barsoum MW. MAX phases: properties of machinable ternary carbides and nitrides. 1st ed. Wiley; 2013. DOI
2. Tzenov NV, Barsoum MW. Synthesis and characterization of Ti_3AlC_2. *J Am Ceram Soc* 2000;83:825-32. DOI
3. Dahlqvist M, Barsoum MW, Rosen J. MAX phases - Past, present, and future. *Mater Today* 2024;72:1-24. DOI
4. Gonzalez-Julian J, Mauer G, Sebold D, Mack DE, Vassen R. Cr_2AlC MAX phase as bond coat for thermal barrier coatings: processing, testing under thermal gradient loading, and future challenges. *J Am Ceram Soc* 2020;103:2362-75. DOI

5. Smialek JL. Oxidation of Al_2O_3 scale-forming MAX phases in turbine environments. *Metall Mater Trans A* 2018;49:782-92. DOI
6. Sun ZM. Progress in research and development on MAX phases: a family of layered ternary compounds. *Int Mater Rev* 2011;56:143-66. DOI
7. Guo L, Yan Z, Wang X, He Q. Ti_2AlC MAX phase for resistance against CMAS attack to thermal barrier coatings. *Ceram Int* 2019;45:7627-34. DOI
8. Lapauw T, Tunca B, Joris J, et al. Interaction of $M_{n+1}AX_n$ phases with oxygen-poor, static and fast-flowing liquid lead-bismuth eutectic. *J Nucl Mater* 2019;520:258-72. DOI
9. Chirica IM, Mirea AG, Neaţu Ş, Florea M, Barsoum MW, Neaţu F. Applications of MAX phases and MXenes as catalysts. *J Mater Chem A* 2021;9:19589-612. DOI
10. Ward L, Dunn A, Faghaninia A, et al. Matminer: an open source toolkit for materials data mining. *Comput Mater Sci* 2018;152:60-9. DOI
11. Sparks TD, Kauwe SK, Parry ME, Tehrani AM, Brgoch J. Machine learning for structural materials. *Annu Rev Mater Res* 2020;50:27-48. DOI
12. Mansouri Tehrani A, Oliynyk AO, Parry M, et al. Machine learning directed search for ultraincompressible, superhard materials. *J Am Chem Soc* 2018;140:9844-53. DOI
13. Sayeed HM, Smallwood W, Baird SG, Sparks TD. NLP meets materials science: quantifying the presentation of materials data in literature. *Matter* 2024;7:723-7. DOI
14. Seegmiller CC, Baird SG, Sayeed HM, Sparks TD. Discovering chemically novel, high-temperature superconductors. *Comput Mater Sci* 2023;228:112358. DOI
15. Alverson M, Baird SG, Murdock R, Ho SH, Johnson J, Sparks TD. Generative adversarial networks and diffusion models in material discovery. *Digit Discov* 2024;3:62-80. DOI
16. Wang AYT, Kauwe SK, Murdock RJ, Sparks TD. Compositionally restricted attention-based network for materials property predictions. *npj Comput Mater* 2021;7:545. DOI
17. Merchant A, Batzner S, Schoenholz SS, Aykol M, Cheon G, Cubuk ED. Scaling deep learning for materials discovery. *Nature* 2023;624:80-5. DOI PubMed PMC
18. Fuhr AS, Sumpter BG. Deep generative models for materials discovery and machine learning-accelerated innovation. *Front Mater* 2022;9:865270. DOI
19. Shetty P, Adeboye A, Gupta S, Zhang C, Ramprasad R. Accelerating materials discovery for polymer solar cells: data-driven insights enabled by natural language processing. arXiv. [Preprint.] Jun 22, 2024 [accessed 2024 Aug 24]. Available from: https://arxiv.org/abs/2402.19462.
20. Alghofaili YA, Alghadeer M, Alsaui AA, Alqahtani SM, Alharbi FH. Accelerating materials discovery through machine learning: predicting crystallographic symmetry groups. *J Phys Chem C* 2023;127:16645-53. DOI
21. Ward L, Agrawal A, Choudhary A, Wolverton C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Comput Mater* 2016;2:16028. DOI
22. Li H, Li X, Li Y, et al. Machine learning assisted design of aluminum-lithium alloy with high specific modulus and specific strength. *Mater Design* 2023;225:111483. DOI
23. Mohanty T, Chandran KSR, Sparks TD. Machine learning guided optimal composition selection of niobium alloys for high temperature applications. *APL Mach Learn* 2023;1:036102. DOI
24. Cover MF, Warschkow O, Bilek MMM, McKenzie DR. A comprehensive survey of M_2AX phase elastic properties. *J Phys Condens Matter* 2009;21:305403. DOI PubMed
25. Head T, Kumar M, Nahrstaedt H, Louppe G, Shcherbatyi I. Scikit-optimize/scikit-optimize. 2020. Available from: https://zenodo.org/records/4014775. [Last accessed on 24 Aug 2024].
26. Ali MA, Hossain MM, Uddin MM, Islam AKMA, Naqib SH. The rise of 212 MAX phase borides: DFT insights into the physical properties of Ti_2PB_2, Zr_2PbB_2, and Nb_2AB_2 [A = P, S] for thermomechanical applications. *ACS Omega* 2023;8:954-68. DOI PubMed PMC
27. Zhang Q, Zhou Y, San X, et al. Zr_2SeB and Hf_2SeB: two new MAB phase compounds with the Cr_2AlC-type MAX phase (211 phase) crystal structures. *J Adv Ceram* 2022;11:1764-76. DOI
28. Rackl T, Johrendt D. The MAX phase borides Zr_2SB and Hf_2SB. *Solid State Sci* 2020;106:106316. DOI
29. Rackl T, Eisenburger L, Niklaus R, Johrendt D. Syntheses and physical properties of the MAX phase boride Nb_2SB and the solid solutions $Nb_2SB_xC_{1-x}$(x = 0-1). *Phys Rev Mater* 2019;3:054001. DOI
30. Ashton M, Hennig RG, Broderick SR, Rajan K, Sinnott SB. Computational discovery of stable M_2AX phases. *Phys Rev B* 2016;94:054116. DOI
31. Ohmer D, Qiang G, Opahle I, Singh HK, Zhang H. High-throughput design of 211 - M_2AX compounds. *Phys Rev Mater* 2019;3:053803. DOI
32. Ohmer D, Opahle I, Singh HK, Zhang H. Stability predictions of magnetic M_2AX compounds. *J Phys Condens Matter* 2019;31:405902. DOI PubMed