# DISCOVERY OF HIGH ENTROPY CERAMICS WITH LOW THERMAL CONDUCTIVITY THROUGH MACHINE LEARNING

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Menglin Wang

AUGUST 2023

© 2023 MENGLIN WANG

**ABSTRACT**

High-entropy ceramics (HEC) materials have gained significant attention for their unique composition and diverse properties. Low thermal conductivity (κ) makes them promising materials for high-temperature applications. However, predicting the thermal properties of HECs is challenging due to the complex nature of their atomic interactions. This project explores the use of machine learning (ML) techniques to discover new compositions of HEC materials with low κ. To achieve this, a dataset combining experimental κ values of 166 HECs and conventional ceramics at various temperatures (5 K-1832 K) is constructed. The input features are then generated based on chemical composition. In the next learning process, various nonlinear regression ML models, including Kernel Ridge Regression (KRR), Random Forest Regression (RF), Support Vector Regression (SVR), and eXtreme Gradient Boosting (XGBoost) are employed. As a result, the RF, KRR, and XGBoost exhibit excellent performance, achieving high $R^2$ scores over 0.90. Additionally, the accuracy of this model was tested using new cases of four compounds, which was not seen for the model before, where a good matching between experimental and predicted κ of the new HECs was found. To gain further insights into the ML models, feature importance analyses are conducted to identify key compositional features that influence κ in HECs. As an application, ML models are leveraged to explore a large composition space and screen promising HECs (30,382) to find candidates with ultralow κ (<1 $W \cdot m^{-1}K^{-1}$) at room temperature. Among the screened HECs, the relationship between important features and the predicted κ is investigated. Size disorder is identified as an important feature in lower κ values of HECs. Overall, this study provides valuable insights into the thermal properties of HECs and demonstrates the efficacy of ML in HECs discovery and design. Accordingly, using the periodic table, compositional information, and ML models, the κ-temperature behavior of a huge number of HECs can be accurately predicted.

## BIOGRAPHICAL SKETCH

Menglin Wang, was born in 1998 in Hebei, China. In 2016, he graduated from Hebei Xinji High School. Following his high school education, he embarked on an academic journey at Anhui University of China, where he pursued and successfully obtained a Bachelor of Engineering degree in New Energy Materials and Devices.

During his undergraduate years, Menglin Wang had the opportunity to engage in extensive research work at Westlake University while completing his thesis. Fascinated by the intersection of Materials Engineering and Computer Science, Menglin decided to pursue a Master of Science degree at Cornell University, focusing on Materials Engineering and Computer Science.

During his time at Cornell, Menglin had the privilege of working under the guidance of Professor Jingjie Yeo and Professor Carla Gomes. Through their mentorship, he delved into the realms of machine learning, cloud computing, and software development, which piqued his interest significantly. Looking ahead, Menglin aspires to build his career in computer science and software development. With a strong foundation in Materials Engineering and a growing expertise in computer science, he aims to leverage his interdisciplinary knowledge to make valuable contributions in these areas.

**ACKNOWLEDGMENTS**

I would like to extend my heartfelt gratitude to Professor Jingjie Yeo, my thesis advisor, for his exceptional support, guidance, and unwavering patience throughout the entire process of my thesis project. Professor Yeo's availability, whether in his office or through online meetings, allowed for insightful discussions and invaluable guidance whenever I faced challenges or encountered obstacles. His constructive feedback and suggestions have played a crucial role in shaping my research project and enhancing my writing. I would also like to express my gratitude to Professor Carla Gomes, my special committee member, for her support.

I am deeply thankful to the staff of the Department of Materials Science and Engineering for their assistance, provision of resources, and continuous encouragement throughout my graduate studies. Special thanks go to Professor Donnelly, the director of the Graduate Study, for her unwavering support and guidance throughout the program.

Finally, I would like to express my profound appreciation to my parents and my girlfriend Jessie for their unconditional love, unwavering support, and constant encouragement throughout my academic journey. Their support during my studies in the United States and their consistent belief in my abilities have been instrumental in my success. Without their love and support, this achievement would not have been possible.

**TABLE OF CONTENTS**

ABSTRACT ..... iii
BIOGRAPHICAL SKETCH ..... iv
ACKNOWLEDGMENTS ..... v
LIST OF FIGURES ..... vii
LIST OF TABLES ..... ix
LIST OF SYMBOLS & ABBREVIATIONS ..... x
INTRODUCTION ..... 1
METHODOLOGY ..... 4
2.1 Data Collection and Characteristics ..... 4
2.2 Feature Engineering ..... 5
2.2.1 Compositional Features ..... 6
2.2.2 Size Disorder ..... 8
2.2.3 Dimensionality Reduction ..... 8
2.3 Machine Learning Models ..... 10
2.3.1 Model Selection ..... 10
2.3.2 Model Training and Testing ..... 11
RESULTS AND DISCUSSION ..... 13
3.1 Model Performance ..... 13
3.1.1 Test Result of Four ML Models ..... 13
3.1.2 Reproduce Experimental Data ..... 16
3.1.3 Impact of Anion Content on κ of HECs ..... 20
3.1.4 Discovery of Promising HECs with Low κ ..... 22
3.2 Feature Importance Analysis ..... 25
3.2.1 Maximum Mendeleev Number ..... 28
3.2.2 Minimum Total Unfilled States ..... 29
3.2.3 Number of Components ..... 30
3.2.4 Size Disorder In Lowering κ of HECs ..... 30
3.3 Limitations ..... 32
CONCLUSION ..... 33
REFERENCES ..... 34
REFERENCES – DATASET ..... 38

## LIST OF FIGURES

Figure 1 Cations (left) and anions (right) distribution in the dataset. ..... 4

Figure 2 Temperature values distribution in the dataset. ..... 5

Figure 3 Pearson correlation map of the features after processing using VT and PC ..... 10

Figure 4 Full κ and its representative split (training and testing) used in the present work. ..... 12

Figure 5 Testing results of the representative split KRR. ..... 14

Figure 6 Testing results of the representative split using RF. ..... 14

Figure 7 Testing results of the representative split XGBoost. ..... 15

Figure 8 Testing results of the representative split SVR. ..... 15

Figure 9 Experimental and ML-predicted temperature-dependent behaviors of $Lu_3TaO_7$ ..... 16

Figure 10 Experimental and ML-predicted temperature-dependent behaviors of Sm3TaO7 ..... 17

Figure 11 Experimental and ML-predicted temperature-dependent behaviors of Dy3TaO7 ..... 17

Figure 12 Experimental and ML-predicted temperature-dependent behaviors of $Ho_{0.5}Lu_{0.5}Yb_{0.5}Eu_{0.5}SiO_5$ ..... 18

Figure 13 Experimental and ML-predicted temperature-dependent behaviors of TiZrHfVTaC ..... 18

Figure 14 Experimental and ML-predicted temperature-dependent behaviors of $Sm_{0.6}Lu_{0.6}Yb_{0.6}Y_{0.6}Dy_{0.6}TaO_7$ ..... 19

Figure 15 RMSE of the Experimental and ML-predicted temperature-dependent κ. ..... 20

Figure 16 Experimental temperature-dependent behavior of (TiZrHfVNbTa)Cx with varied carbon content ..... 21

Figure 17 KRR predicted temperature-dependent behavior of (TiZrHfVNbTa)Cx with varied carbon content ..... 22

Figure 18 Selected cations and anions for searching new HECs. ..... 23

Figure 19 Feature importance ranking of RF. ..... 26

Figure 20 Feature importance ranking of XGBoost. ..... 27

Figure 21 Plot of XGBoost predicted κ (averaged) as function to Maximum Mendeleev Number.....28

Figure 22 Plot of XGBoost predicted κ (averaged) as function to Minimum Total Unfilled States.....29

Figure 23 Plot of XGBoost predicted κ (averaged) as function to Number of Components.....30

Figure 24 Scatter of XGBoost predicted κ (averaged) with Size Disorder on RT-ultralow HECs.....31

## LIST OF TABLES

Table 1 22 elemental properties used for the feature calculation. .....6
Table 2 Features and their abbreviation selected to train ML models. .....9
Table 3 Test Performance of four algorithms. .....13
Table 4 XGBoost predicted HECs with RT-ultralow κ. .....24
Table 5 XGBoost predicted HECs with RT-extremely low κ. .....25

## LIST OF SYMBOLS & ABBREVIATIONS

| Symbol | Remarks |
|---|---|
| % | Percentage |
| $\kappa$ | Thermal conductivity |
| $\kappa_L$ | Lattice thermal conductivity |
| $\kappa_e$ | Electron thermal conductivity |
| $\delta$ | Size disorder |
| K | Kelvin, SI unit of temperature |
| $\sum$ | Sum of numbers |
| f | Property of element |
| x | Atomic fraction |
| I | Fraction of ionic character |
| X | Electronegativity |
| $R^2$ | Coefficient of determination |
| T | Temperature |
| m | Meter, SI unit of length |
| W | Watt, SI unit of power |
| $\sigma$ | Electronic conductivity |

| Abbreviation | Remarks |
|---|---|
| HEC | High Entropy Ceramics |
| ML | Machine Learning |
| KRR | Kernel Ridge Regression |
| RF | Random Forest |
| XGBoost | eXtreme Gradient Boosting |
| SVR | Support Vector Regression |
| VT | Variance Threshold |
| PC | Pearson Correlation |
| GridSearchCV | Grid Search Cross-Validation |
| RT | Room Temperature |
| MAE | Mean Absolute Error |
| RMSE | Root Mean Square Error |

# CHAPTER 1

# INTRODUCTION

High entropy ceramics (HECs) have garnered significant attention since 2015 and are regarded as a promising category of materials with diverse applications and distinctive properties (Oses et al., 2020). These materials exhibit a unique composition, typically comprising at least four different cations or anions within a single phase. This composition endows HECs with a broad range of properties. Researchers have successfully synthesized and investigated various types of HEC materials, including nitrides, borides, carbides, and oxides, to explore their properties (Gild et al., 2016; Albedwawi et al., 2021; Liu et al., 2020; Moskovskikh et al., 2020). One notable property of HECs is their low thermal conductivity (κ), which is crucial for applications in high-temperature components of gas turbines, aero engines, automobiles, and other systems. By utilizing ceramics with low κ, it is possible to enhance the efficiency of these components by enabling their operation at higher temperatures (Padture, 2002; Clarke et al, 2005).

In recent years, significant advancements have been made in the calculation of κ of HECs using molecular dynamics. A study conducted by A. Giri et al. utilized the widely employed 12-6 Lennard Jones (LJ) potential and the non-equilibrium molecular dynamics (NEMD) method to examine the lattice thermal conductivity of a five-component solid solution (A. Giri et al., 2018). In addition, M. Lim et al. employed the equilibrium molecular dynamics (EMD) method to investigate the thermal conductivity of J14 $(Mg_{0.2}Co_{0.2}Ni_{0.2}Cu_{0.2}Zn_{0.2})O$ (M. Lim et al., 2019). However, it is important to note that the interatomic potentials employed in these models are oversimplified compared to real systems, as it is challenging to develop highly reliable potentials for complex high-entropy systems. Density Functional Theory (DFT) provides a comprehensive framework for modeling intricate atomic interactions. As a parameter-free method, DFT can provide input parameters for other formalisms used in lattice thermal conductivity calculations. However, due to its high computational cost, DFT has not yet been used for high entropy systems, while its application for binary solid solution has given some promising results (Zhang et al., 2019).

Over the past three years, there has been a significant focus on utilizing machine learning (ML) techniques in the field of materials discovery and design. These techniques have allowed researchers to circumvent the complex and costly conventional processes involved in designing new materials for various applications. This is largely due to the significant advancements in ML algorithms and the availability of powerful computing resources. Several researchers have applied ML models to investigate the thermal conductivity of compounds. For example, Seko et al. used a kernel ridge regression model on a dataset of 110 compounds with lattice thermal conductivity ($\kappa_L$) computed by the DFT method (Seko et al., 2017). Chen et al. developed a machine learning model based on a Gaussian process regression algorithm, using a dataset of 100 inorganic materials with experimentally measured $\kappa_L$ (Chen et al., 2019). Carrete et al utilized a random forest regression model trained on a dataset of 32 half-Heusler compounds, calculated using ab initio methods. (Carrete et al., 2014). However, there have been limited studies on predicting the thermal properties of HECs, despite their unique composition and properties.

In this study, we investigate the discovery of new compositions of high-entropy ceramic (HEC) materials with low $\kappa$ using various nonlinear regression machine learning (ML) models. The ML models we employ include Kernel Ridge Regression (KRR), Random Forest Regression (RF), Support Vector Regression (SVR), and eXtreme Gradient Boosting (XGBoost). To accomplish this, we compile a dataset that combines $\kappa$ values from both HEC materials and conventional ceramic materials. Our focus lies in utilizing compositional features derived from four attributes as inputs for the ML models.

Among the models tested, RF, KRR, and XGBoost demonstrate excellent performance in predicting $\kappa$, achieving R2 scores of 0.91, 0.91, and 0.90, respectively. The mean square errors (MAE) for these models are 1.34 $W{\cdot}m^{-1}K^{-1}$, 1.28 $W{\cdot}m^{-1}K^{-1}$ and 1.10 $W{\cdot}m^{-1}K^{-1}$, respectively. Furthermore, we conduct a feature importance analysis using RF and XGBoost models to gain insights into the underlying mechanisms that influence thermal conductivity in HECs. These analyses allow us to identify significant composition features that affect thermal conductivity in HECs. As an application of the ML models developed, we utilize

this framework to search for new HECs with low lattice thermal conductivity among 30,382 generated compositions. In summary, the study achieves high prediction accuracy in predicting the κ-temperature behavior of HECs and provides valuable insights into the composition features that influence the thermal conductivity in HECs.

## CHAPTER 2

## METHODOLOGY

### 2.1 Data Collection and Characteristics

The initial dataset used in this study consists of κ values of various ceramic materials, including both conventional ceramics and high-entropy ceramics (HECs). The dataset is constructed based on previous experimental works. Each composition in the dataset comes from literature that listed in REFERENCES – DATASET and we include the entire dataset in the supplementary files. The initial dataset comprises thermal conductivity data for 166 compositions. Among these compositions, there are 93 HECs and 73 conventional ceramics. The κ values are measured at various temperatures ranging from 0.14 $W \cdot m^{-1}K^{-1}$ to 100 $W \cdot m^{-1}K^{-1}$ and most of the values fall below 3 $W \cdot m^{-1}K^{-1}$. Regarding the cation elements present in the compositions, the majority of cases involve Zr (259 cases), along with other elements such as Nb (101 cases) and Y (96 cases). As for the distribution of anion elements, the dataset contains cases with O (418 cases) and C (125 cases) (Figure 1).

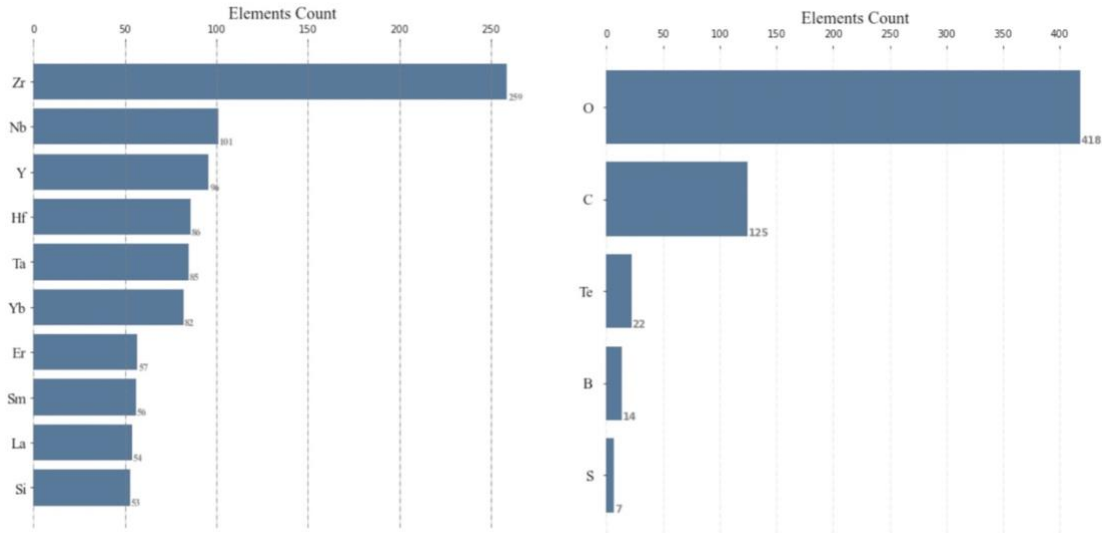


*Figure 1 Cations (left) and anions (right) distribution in the dataset.*

The temperature range of the measured κ values is wide, spanning from 5 K to 1832 K. Moreover, the distribution of temperature measurements indicates a concentration of around 300 K (Figure 2). This distribution provides a suitable basis for training the machine learning model. Such a model can effectively predict thermal conductivity for a wide range of compositions and capture temperature-dependent behavior. Additionally, considering the range κ values in the dataset, the model can identify compositions that are expected to exhibit low thermal conductivity.

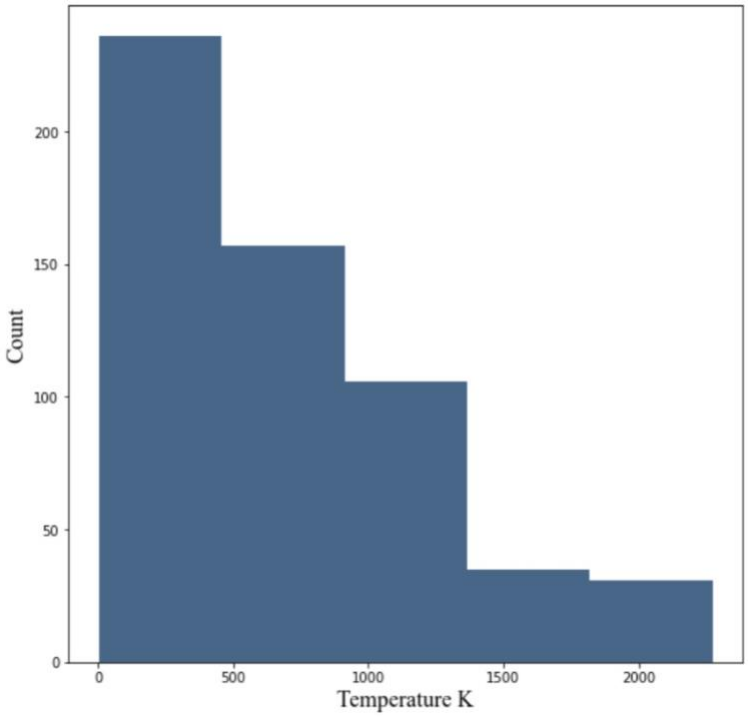


*Figure 2 Temperature values distribution in the dataset.*

### 2.2 Feature engineering

Prior to training the machine learning (ML) models, compositional features were generated based on four attributes, including elemental-property-based attributes, stoichiometric attributes, valance orbital occupation attributes, and ionic compound attributes (Ward et al., 2016).

It should be noted that crystal structure features and processing parameters are not included in this work.

Because the inclusion of crystal structure features such as formation energy, bandgap, density of states (DOS), and cell volume would necessitate uniform input parameters and a well-defined crystal structure. These parameters are typically obtained through first-principles calculations, which are less commonly available in the literature. Consequently, incorporating such crystal features would result in a significant reduction in the number of cases within the initial dataset. Moreover, since many reported ceramic materials exist in multiphase systems, they cannot be described by a singular crystal structure. By focusing solely on composition features, our machine learning (ML) models can learn from a wider range of compositions, κ values, and temperatures, enabling the development of a more generalized model suitable for high-throughput HEC screening. Processing parameters were not considered in this work due to the lack of uniform reporting in the field of HECs (Arbuzova et al., 2006). Incorporating these parameters would have further reduced the size of the initial dataset.

### 2.2.1 Compositional Features

The elemental-property-based attributes are calculated based on statistics of the 22 elemental properties listed in Table 1.

| | | | |
|---|---|---|---|
| Atomic number | Mendeleev number | s valance electrons | Unfilled s states |
| Atomic weight | Column | p valance electrons | Unfilled p states |
| Melting temperature | Row | d valance electrons | Unfilled d states |
| Electronegativity | Covalent radius | f valance electrons | Unfilled f states |
| Band gap of ground state | Specific volume of ground state | Total valance electrons | Total unfilled states |
| Space group number of ground state | Magnetic moment of ground state | | |

*Table 1 22 elemental properties used for the feature calculation.*

These elemental properties were then processed using six equations: mean, mean absolute deviation, range, maximum, minimum, and mode. The mean and average deviation are calculated using the following formulae:

$$\bar{f} = \sum x_i f_i \tag{1}$$

$$\hat{f} = \sum x_i \left| f_i - \bar{f} \right| \tag{2}$$

where $f_i$ is the property of element i, $x_i$ is the atomic fraction, $\bar{f}$ is the mean, and $\hat{f}$ is the average deviation.

The stoichiometric attributes capture the fraction of the elements present and are not affected by what those elements are. All are based on $L^p$ norms of a vector representing the atomic:

$$\|x\|_p = \left( \sum_{i=0}^{n} |x_i|^p \right)^{\frac{1}{p}} \tag{3}$$

The valance orbital occupation attributes are the fraction-weighted average of the number of valance electrons in each orbital divided by the fraction-weighted average of the total number of valance electrons. This attribute is exactly equivalent to the one employed by Agrawal et al.

As an example, the fraction of p electrons for $Fe_2O_3$ is computed by:

$$F_p = \frac{\frac{2}{5}(0) + \frac{3}{5}(4)}{\frac{2}{5}(8) + \frac{3}{5}(6)} = \frac{6}{17} \cong 0.352$$

The ionic compound attributes are designed to determine whether a material is ionically bonded. The first measure is a Boolean denoting whether it possible to form a neutral, ionic compound assuming each element takes exactly one of its common charge states. The other two are based on the "ionic character" of a binary compound, which is computed from the electronegativity difference between its two constituent elements using the relation:

$$I(X_A, X_B) = 1 - \exp(-0.25(X_A - X_B)^2) \tag{4}$$

where I is the fraction of ionic character, $X_A$ is the electronegativity of element A, and $X_B$ is the

electronegativity of element B. The first attribute we used is the maximum ionic character between any two elements in the material. The second is the mean ionic character, which is computed using:

$$\bar{I} = \sum x_i \, x_j * I(X_i, X_j) \tag{5}$$

### 2.2.2 Size Disorder

In our approach, we took into account the size disorder of the components within the feature space, which are defined as:

$$\delta = \sum_{i=1}^{n} x_i \left(\frac{r_i}{r_{avg}} - 1\right)^2 \tag{6}$$

where $x_i$ is the atomic fraction, $r_i$ is the radius of the $i^{th}$ component, and $r_{avg}$ is the overall weighted average of radius. Here, we used the Shannon effective ionic radius (Shannon, 1976). It is assumed that all cations share the same lattice site, and the most common charge state for each element is considered. By incorporating size disorder into our analysis, we aim to enhance the prediction capabilities specifically for high-entropy ceramics (HECs), and this results in a set of features (132 + 13 + 1 + 1) that was used in the further learning steps.

### 2.2.3 Dimensionality Reduction

To reduce the number of features in the training process, and thus, to reduce the complexity of the final model, the full feature set was processed using variance threshold (VT) and Pearson correlation (PC) statistical methods, respectively. In the variance threshold method, the features that show a small variation between the cases were removed, and this was determined using the variance, where the features with a variance less than 0.10 were considered to be constant and they were supposed to have no effect on the property. The remaining features were further included in a selection process using Pearson correlation, and here, features that are highly correlated (correlation coefficient > 0.90) were excluded. A set of 51

features (Table 2) was finally obtained after the VT and PC processing, the PC heatmap of these features is shown in Figure 3.

| Abbreviation | Feature | Abbreviation | Feature |
|---|---|---|---|
| F1 | Temperature | F27 | most_NfValence |
| F2 | NComp | F28 | mean_NValance |
| F3 | mean_Number | F29 | maxdiff_NValance |
| F4 | maxdiff_Number | F30 | min_NValance |
| F5 | dev_Number | F31 | most_NValance |
| F6 | min_Number | F32 | mean_NpUnfilled |
| F7 | most_Number | F33 | maxdiff_NpUnfilled |
| F8 | mean_MendeleevNumber | F34 | most_NpUnfilled |
| F9 | maxdiff_MendeleevNumber | F35 | mean_NdUnfilled |
| F10 | dev_MendeleevNumber | F36 | maxdiff_NdUnfilled |
| F11 | max_MendeleevNumber | F37 | dev_NdUnfilled |
| F12 | most_MendeleevNumber | F38 | mean_NfUnfilled |
| F13 | maxdiff_MeltingT | F39 | mean_NUnfilled |
| F14 | dev_MeltingT | F40 | maxdiff_NUnfilled |
| F15 | max_MeltingT | F41 | dev_NUnfilled |
| F16 | maxdiff_Column | F42 | min_NUnfilled |
| F17 | dev_Column | F43 | mean_GSvolume_pa |
| F18 | maxdiff_Row | F44 | maxdiff_GSvolume_pa |
| F19 | dev_Row | F45 | dev_GSvolume_pa |
| F20 | mean_Electronegativity | F46 | min_GSvolume_pa |
| F21 | maxdiff_NsValence | F47 | most_GSvolume_pa |
| F22 | dev_NpValence | F48 | most_GSbandgap |
| F23 | maxdiff_NdValence | F49 | max_SpaceGroupNumber |
| F24 | dev_NdValence | F50 | CanFormIonic |
| F25 | max_NdValence | F51 | SizeDisorder |
| F26 | mean_NfValence | | |

*Table 2 Features and their abbreviation selected to train ML models.*

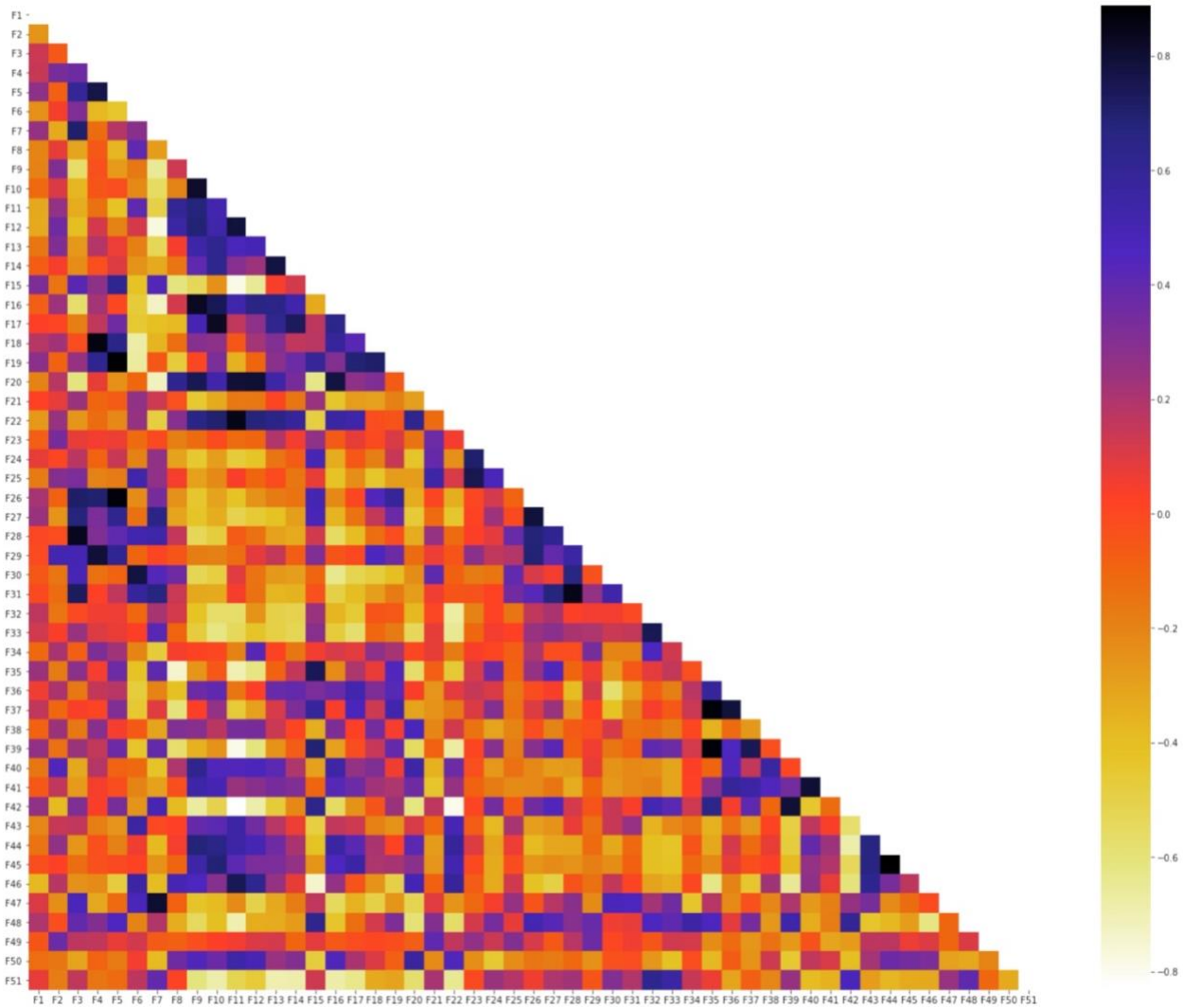


*Figure 3 Pearson correlation map of the features after processing using VT and PC*

## 2.3 Machine learning models.

### 2.3.1 Model Selection

Learning the thermal conductivity behavior for compounds as a function of the temperature was done using different ML algorithms, including Kernel Ridge Regression (KRR), Random Forest Regression (RF), Support Vector Regression (SVR) and eXtreme Gradient Boosting (XGBoost).

KRR is an extension of ridge regression that introduces a mapping function $\Phi^K$ to transform the data from a nonlinear space to a linear space. The kernel K ($x_i$, $x_j$) is defined as the inner product of the mapped data points. By employing the kernel function, KRR can capture nonlinear relationships between the input and output variables. It constructs a hyperplane or set of hyperplanes in a high-dimensional space to map the input data points and predict the corresponding output labels. SVR is particularly useful in dealing with regression problems and can handle complex data relationships. The form of the model learned by KRR is similar to SVR. However, the loss functions are different. RF is an ensemble learning method that combines multiple decision trees to make predictions. Each decision tree in the random forest is trained on a subset of the data, and the final prediction is obtained by averaging the predictions of all the individual trees. RF is known for its ability to handle high-dimensional data and capture complex interactions between features. XGBoost is based on gradient boosting models and works by creating an ensemble of decision trees. The model predicts the output by taking a weighted sum of predictions from multiple models. XGBoost offers better generalization performance compared to other algorithms due to its ability to penalize complexity and utilize computational tricks.

### 2.3.2 Model Training and Testing

The XGBoost regression model is implemented using the XGBoost library in Python, while the other three models (KRR, RF and SVR) are implemented using Scikit-Learn. The ML models were then built through training the dataset containing the target property and the preprocessed features, and for comparison, this process was conducted using the same representative split (training; 80% and testing; 20%) (Figure 4). It is seen that the training subset in these split covers the full range of κ values; this, in turn, results in a ML model that can be used to predict κ in the same range.

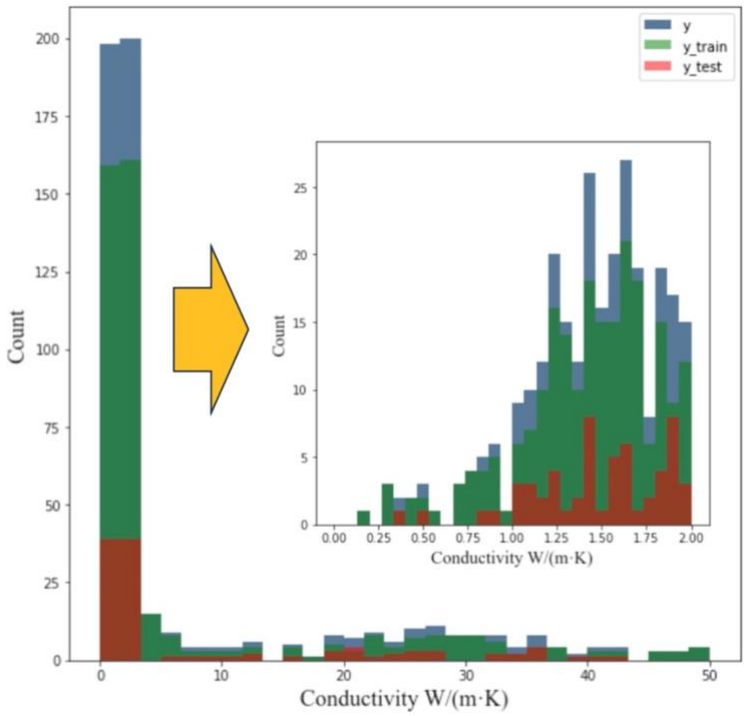


*Figure 4 Full κ and its representative split (training and testing) used in the present work.*

For the training process, the four ML algorithms are trained by the GridSearch method with 5-fold cross validation and grid search to determine the optimal values. GridSearchCV will loop through predefined hyperparameters in a cross-validation way and fit the estimator on your training set and it can select the parameters from the listed hyperparameters which led to the best model performance. K-Fold cross validation is an algorithm that randomly splits the data set into K parts and trains a model on the K-1 parts and tests over the remaining one. This process is repeated for each of the K partitions as test sets, and the average performance is computed to assess the predictive ability of the model.

# CHAPTER 3

# RESULTS AND DISCUSSION

## 3.1 Model Performance

### 3.1.1 Test Result of Four ML Models

Table 3 summarizes the performance of four ML algorithms with the metric of MAE, RMSE and $R^2$ score on the test set and training set. The KRR, RF, and XGBoost ML models demonstrate robust performance, outperforming SVR.

| Model | train result | | | test result | | |
|---|---|---|---|---|---|---|
| | MAE | RMSE | $R^2$ | MAE | RMSE | $R^2$ |
| KRR | 0.306 | 1.015 | 0.990 | 1.104 | 3.920 | 0.903 |
| RF | 0.880 | 2.410 | 0.975 | 1.370 | 3.790 | 0.910 |
| XGBoost | 0.537 | 1.059 | 0.990 | 1.286 | 3.780 | 0.910 |
| SVR | 3.139 | 9.340 | 0.625 | 2.511 | 5.400 | 0.816 |

*Table 3 Test Performance of four algorithms.*

The testing results of the four models are shown in Figure 5-8, respectively. The RF model achieves an R2 score on the test set of 0.91 with an MAE score of 1.33 $W \cdot m^{-1}K^{-1}$. Similarly, the XGBoost model achieves a high R2 score of 0.91 and an MAE score of 1.28 $W \cdot m^{-1}K^{-1}$. The KRR model also showcases the strong performance, attaining an R2 score on the test set of 0.90 and an MAE score of 1.1 $W \cdot m^{-1}K^{-1}$.

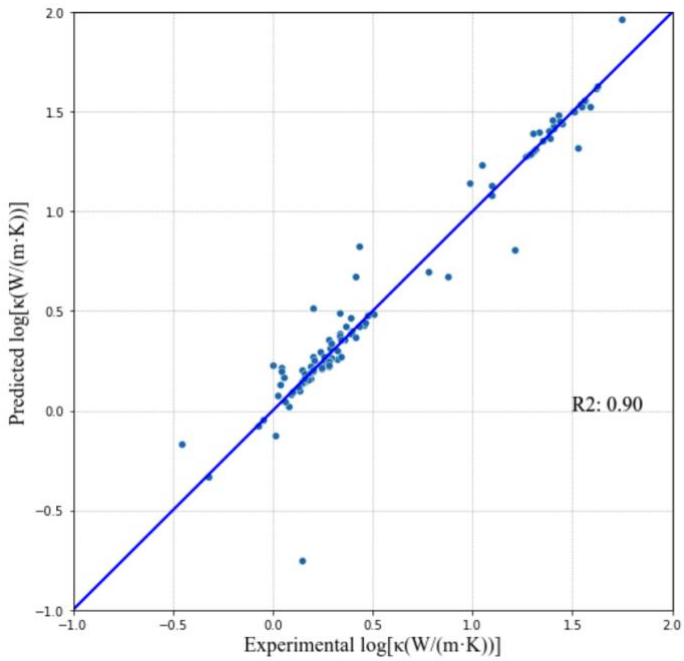


*Figure 5 Testing results of the representative split KRR.*

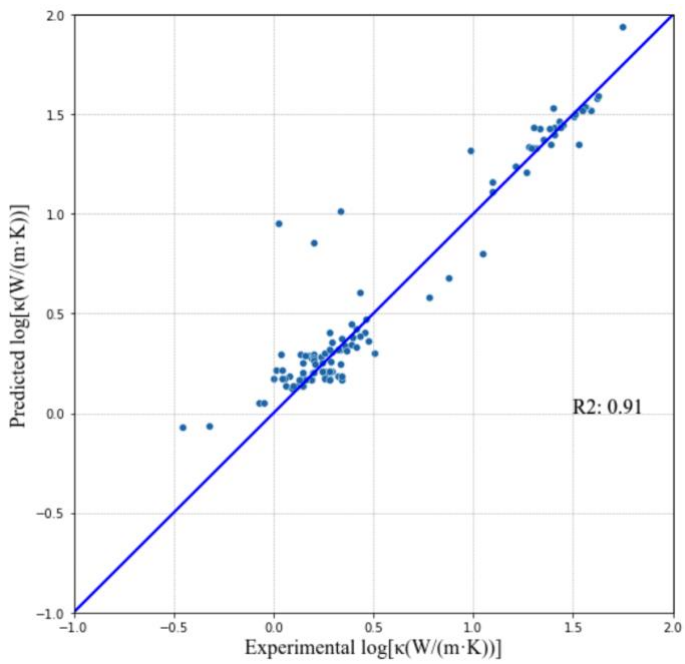


*Figure 6 Testing results of the representative split using RF.*

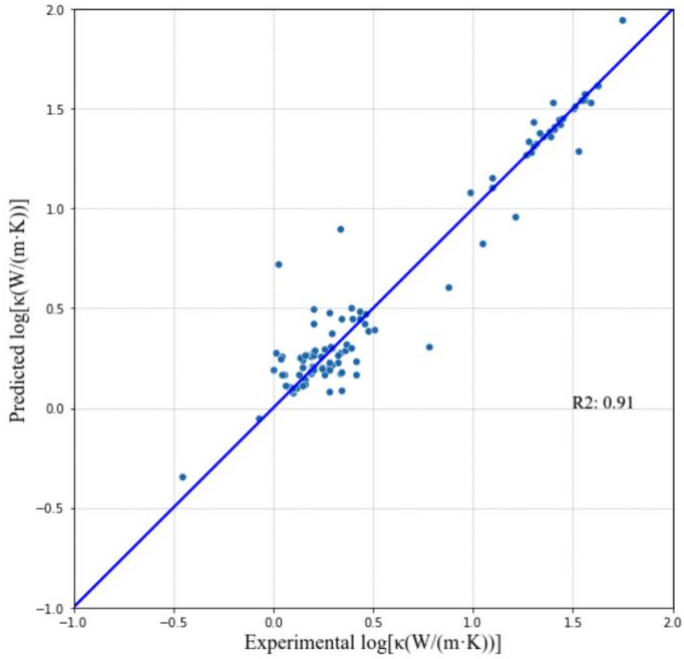


*Figure 8 Testing results of the representative split XGBoost.*

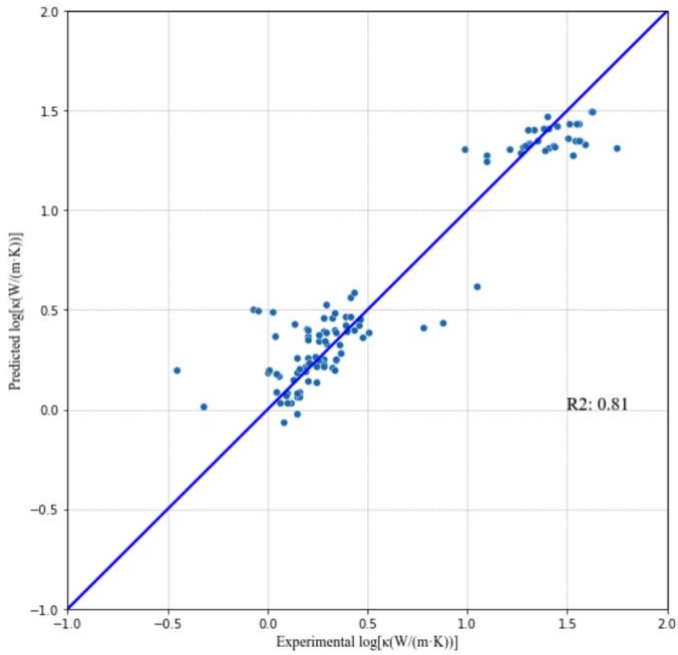


*Figure 7 Testing results of the representative split SVR.*

### 3.1.2 Reproduce Experimental Data

The ability of the ML models to reproduce experimental data was examined by comparing the predicted κ with the measured values of several ceramics and HECs. Specifically, the ceramics $Lu_3TaO_7$, $Sm_3TaO_7$, $Dy_3TaO_7$ and HECs TiZrHfVNbTaC, $Ho_{0.5}Lu_{0.5}Yb_{0.5}Eu_{0.5}SiO_5$, $Sm_{0.6}Lu_{0.6}Yb_{0.6}Y_{0.6}Dy_{0.6}TaO_7$ (Sang et al., 2023; Chen et al., 2022; Chen et al., 2023) conducted recently were considered for comparison. According to the results shown in Figure 9-14, one can find that the experimental data is reproducible using the three models for all cases.

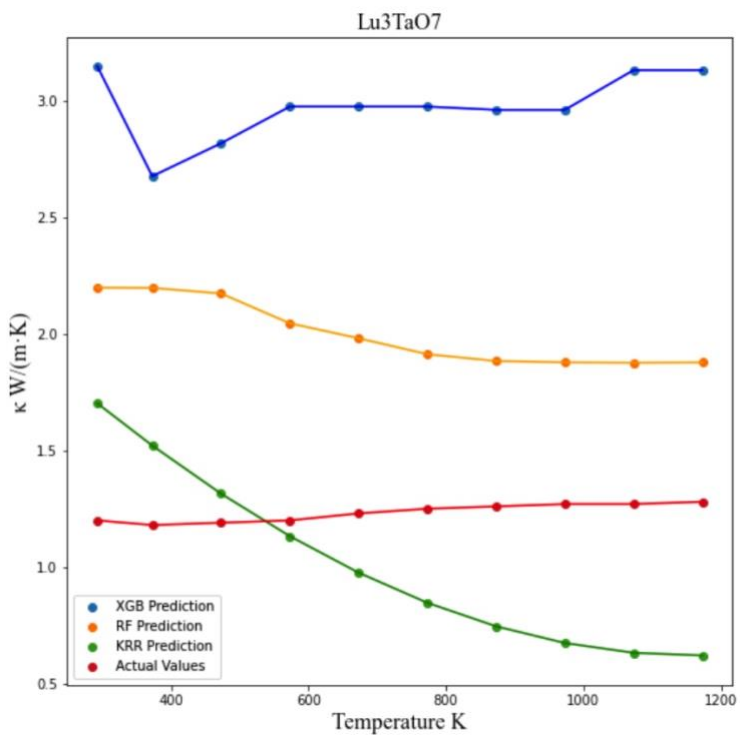


*Figure 9 Experimental and ML-predicted temperature-dependent behaviors of $Lu_3TaO_7$.*

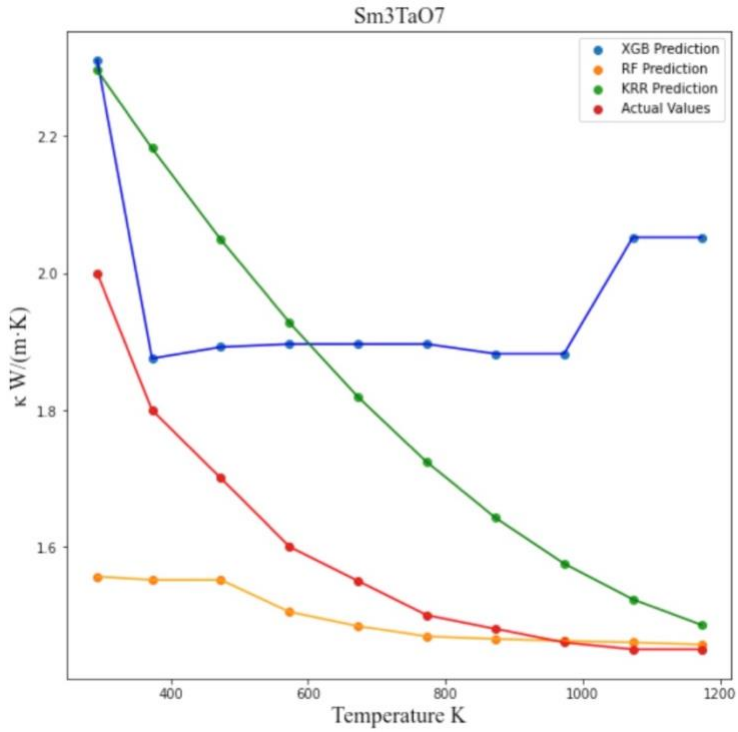


*Figure 10 Experimental and ML-predicted temperature-dependent behaviors of Sm3TaO7*

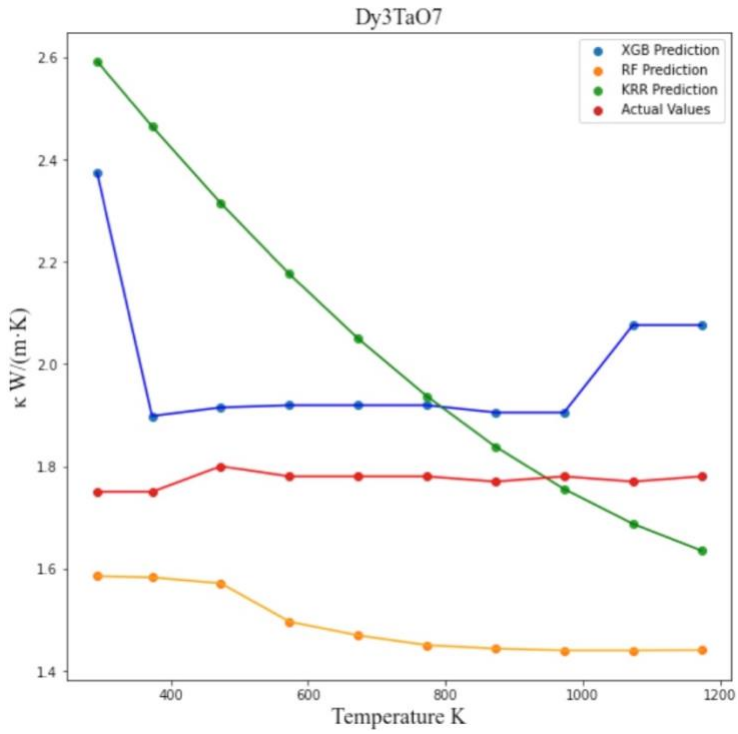


*Figure 11 Experimental and ML-predicted temperature-dependent behaviors of Dy3TaO7.*

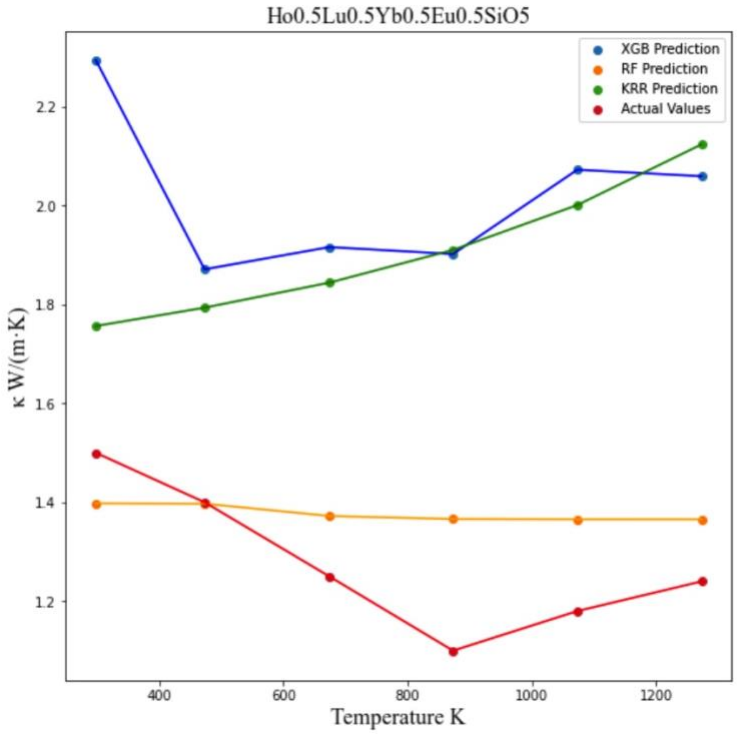


*Figure 13 Experimental and ML-predicted temperature-dependent behaviors of $Ho_{0.5}Lu_{0.5}Yb_{0.5}Eu_{0.5}SiO_5$*

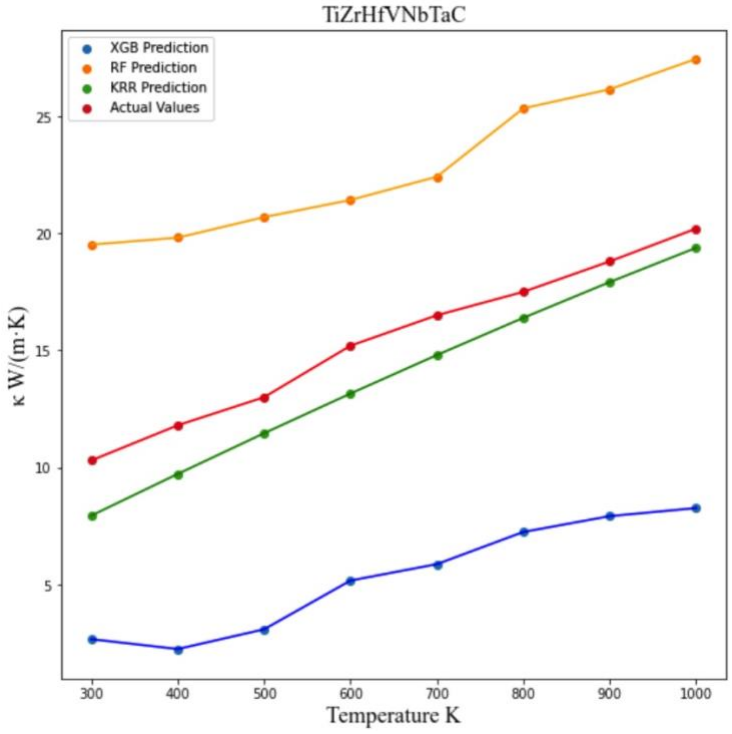


*Figure 12 Experimental and ML-predicted temperature-dependent behaviors of TiZrHfVTaC*

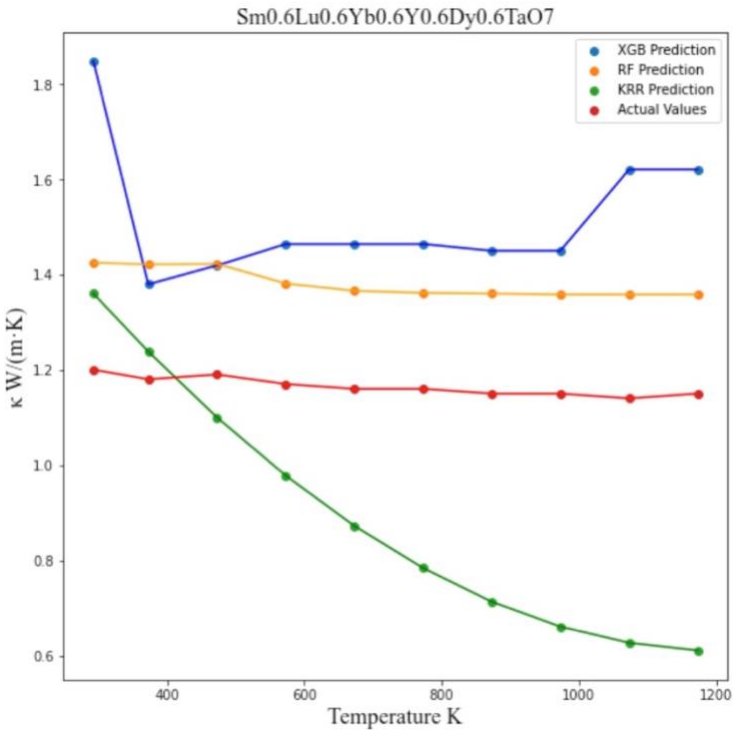


*Figure 14 Experimental and ML-predicted temperature-dependent behaviors of $Sm_{0.6}Lu_{0.6}Yb_{0.6}Y_{0.6}Dy_{0.6}TaO_7$*

.

Additionally, the root-mean-square error (RMSE) values, calculated between the experimental and predicted κ values, are summarized in Figure 15. Overall, these findings indicate that the ML models are capable of accurately predicting thermal conductivity and can successfully replicate experimental data. While there may be variations in the performance of the models for different compounds, they generally exhibit a good understanding of the temperature-dependent behavior of thermal conductivity.

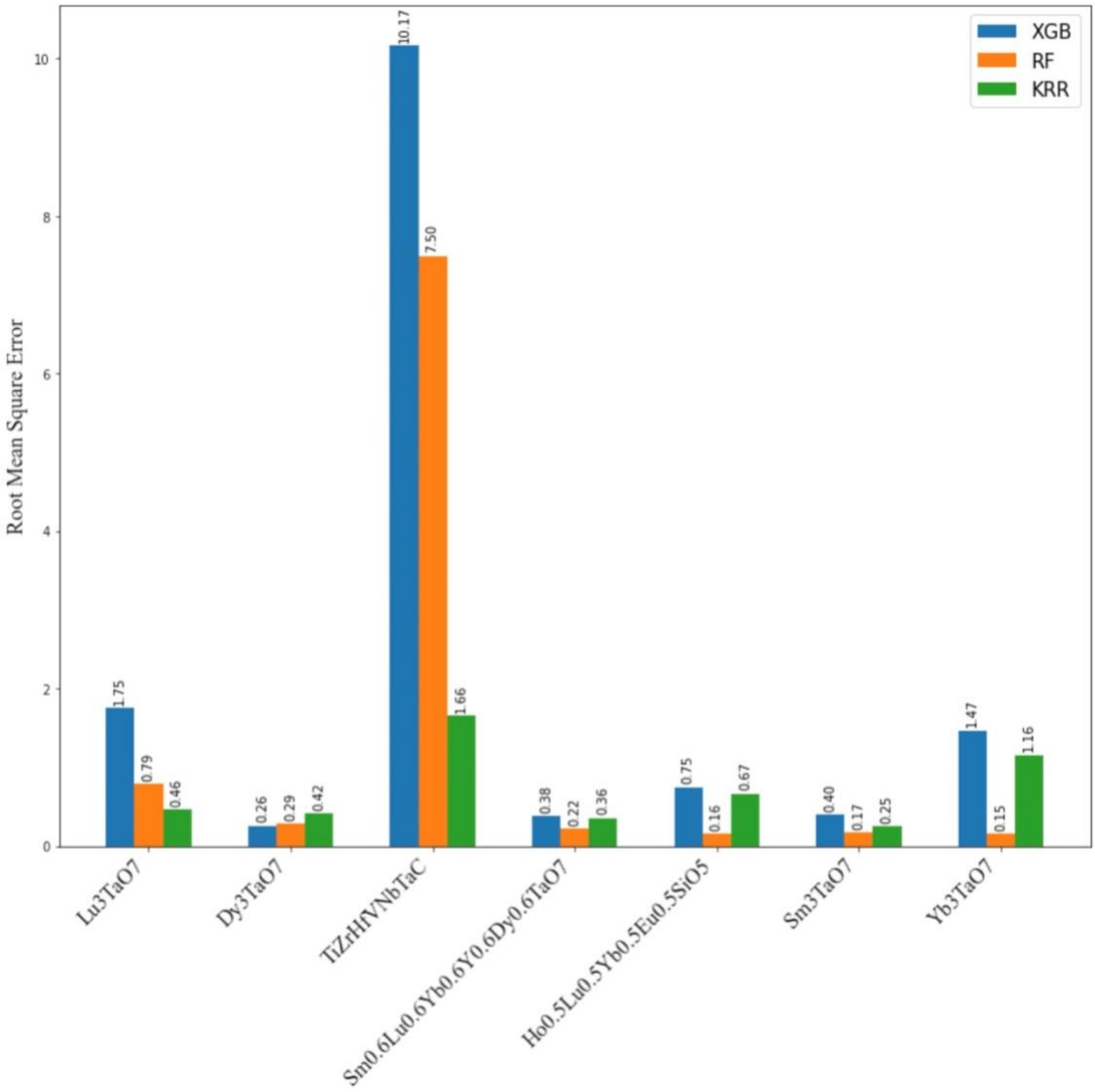


*Figure 15 RMSE of the Experimental and ML-predicted temperature-dependent κ.*

### 3.1.3 Impact of anion content on κ of HECs

In our investigation, we also examined the impact of the anion content on the thermal conductivity of HECs using our ML model. Specifically, we focused on the (TiZrHfVNbTa)Cx (x = 0.6 – 1) compounds. Chen et al. revealed that the thermal conductivity decreases as the carbon content decreases (Figure 16). This behavior is attributed to the increased scattering of conducting phonons and electrons caused by the presence of carbon vacancies (Chen et al., 2023).

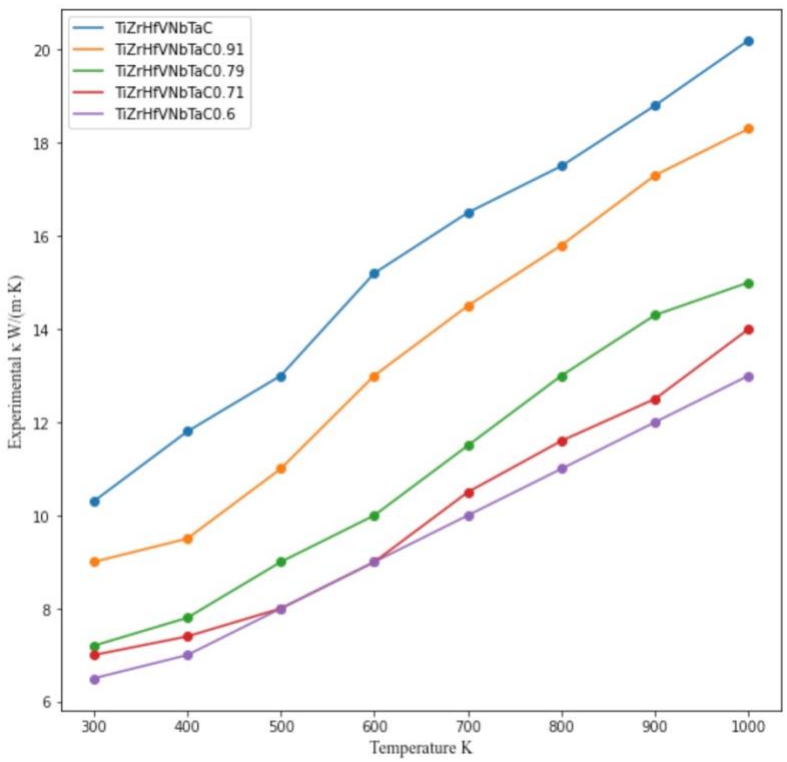


*Figure 16 Experimental temperature-dependent behavior of (TiZrHfVNbTa)Cx with varied carbon content.*

To assess the performance of our ML model in predicting (TiZrHfVNbTa)Cx compounds with varying carbon content, we selected the KRR model, which exhibited a smaller prediction error for TiZrHfVNbTaC. Figure 17 illustrates the results obtained using the KRR model. Notably, the model successfully captures the trend of increasing thermal conductivity with elevated temperatures. Furthermore, it reproduces the observed trend of decreasing thermal conductivity with decreasing carbon content, demonstrating the robust

capability of the KRR model in predicting HECs with varied anion content.

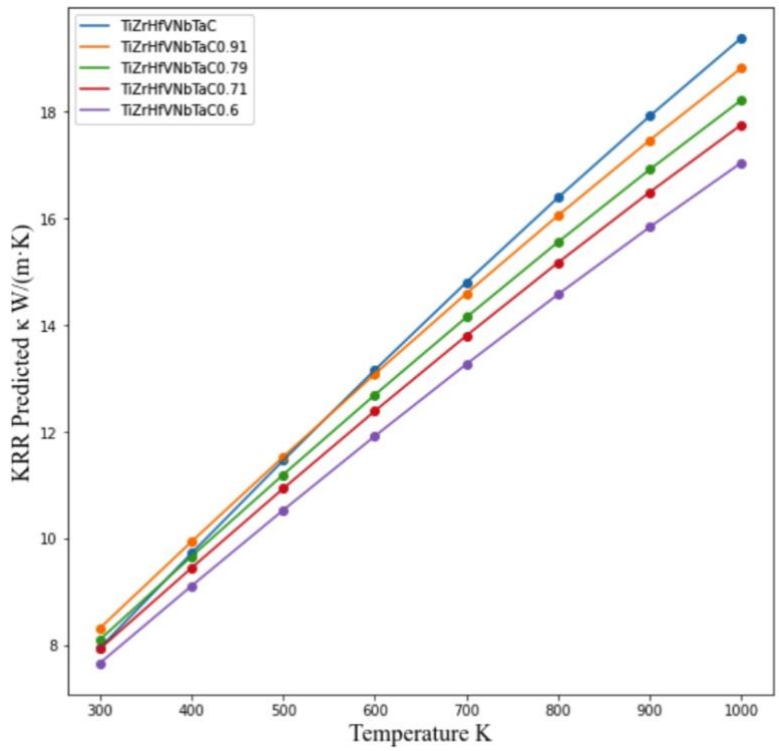


*Figure 17 KRR predicted temperature-dependent behavior of (TiZrHfVNbTa)Cx with varied carbon content.*

### 3.1.4 Discovery of Promising HECs with Low κ

In order to identify high-entropy ceramics (HECs) with low thermal conductivity and extremely low thermal conductivity, we utilized the predictive XGBoost model. We generated features for a total of 30,382 HECs, following the same methodology as before, ensuring that each compound consisted of 2 to 6 cations and one anion. The selected cations and anions are highlighted in Figure 18, indicating their positions in the periodic table.

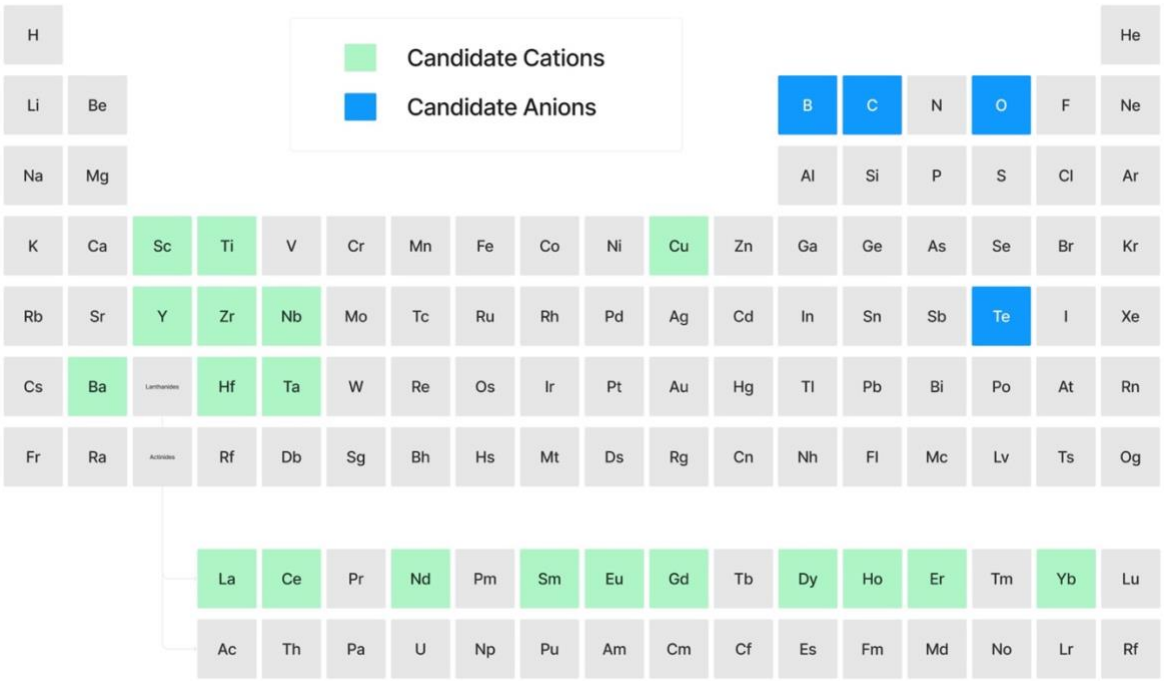


*Figure 18 Selected cations and anions for searching new HECs.*

To further refine the pool of candidate materials, we employed a screening process based on solid solubility. In line with the Hume-Rothery rules and considering the size effects of component atoms, we utilized an ionic radius criterion ($\delta < 1$) to evaluate the compatibility of different elements. As a result, 24,179 compounds were identified as potential ceramics with the potential for low thermal conductivity. The generated features of these candidate compounds were then fed into the XGBoost model, which predicted their thermal conductivity values at room temperature (RT). Based on these predictions, we identified 246 compounds with RT-ultralow thermal conductivity and 9 compounds with RT-extremely low thermal conductivity, and they are listed in Table 4 and Table 5 respectively.

| | | | |
|---|---|---|---|
| $ZrNbHfCuO_{10}$ | $YbErLaDyTe_{10}$ | $YbErSmGdLaTe_{10}$ | $NbHfLaEuGdScO_{10}$ |
| $ZrHfTaCuO_{10}$ | $YbErLaGdTe_{10}$ | $YbErSmLaNdTe_{10}$ | $NbHfDyEuBaTiO_{10}$ |
| $ZrHfTiScO_{10}$ | $YbErDyLaTe_{10}$ | $YbErLaDyGdTe_{10}$ | $NbHfDyEuBaScO_{10}$ |
| $ZrGdTiScO_{10}$ | $YbErDyNdTe_{10}$ | $YbErLaDyNdTe_{10}$ | $NbHfDyEuLaScO_{10}$ |
| $NbHfTiScO_{10}$ | $YbErGdLaTe_{10}$ | $YbErDyGdLaTe_{10}$ | $NbHfDyGdBaScO_{10}$ |
| $NbHfScCuO_{10}$ | $YbSmDyGdTe_{10}$ | $YbErDyLaNdTe_{10}$ | $NbHfDyBaCeScO_{10}$ |
| $NbErSmDyTe_{10}$ | $YbSmDyCuTe_{10}$ | $YbSmLaDyGdTe_{10}$ | $NbHfEuGdBaScO_{10}$ |
| $NbGdTiScO_{10}$ | $YbSmBaNdTe_{10}$ | $YbSmLaNdCuTe_{10}$ | $NbHfEuGdLaScO_{10}$ |
| $NbCeTiScO_{10}$ | $ErSmDyGdTe_{10}$ | $YbSmDyGdLaTe_{10}$ | $YErSmLaDyGdTe_{10}$ |
| $YHfTiScO_{10}$ | $ZrNbYTiScO_{10}$ | $YbSmEuBaNdTe_{10}$ | $YErSmLaDyNdTe_{10}$ |
| $YErSmLaTe_{10}$ | $YErSmLaDyTe_{10}$ | $YbSmLaNdCuTe_{10}$ | $YErSmDyGdLaTe_{10}$ |
| $YErSmLaTe_{10}$ | $YErSmLaGdTe_{10}$ | $ZrNbHfSmDyBaO_{10}$ | $YErSmDyLaNdTe_{10}$ |
| $YErSmNdTe_{10}$ | $YErSmLaNdTe_{10}$ | $ZrYLaDyEuGdTe_{10}$ | $YSmLaDyEuGdTe_{10}$ |
| $YErLaDyTe_{10}$ | $YErSmDyLaTe_{10}$ | $ZrYDyEuGdLaTe_{10}$ | $YSmLaDyEuNdTe_{10}$ |
| $YErLaGdTe_{10}$ | $YErSmDyNdTe_{10}$ | $NbYHfErSmBaO_{10}$ | $YSmDyEuGdLaTe_{10}$ |
| $YErLaNdTe_{10}$ | $YErSmGdLaTe_{10}$ | $NbYHfErSmCeO_{10}$ | $YSmDyEuLaNdTe_{10}$ |
| $YErDyLaTe_{10}$ | $YErSmLaNdTe_{10}$ | $NbYHfErDyBaO_{10}$ | $YDyEuBaNdScTe_{10}$ |
| $YErDyNdTe_{10}$ | $YErLaDyGdTe_{10}$ | $NbYHfErDyCeO_{10}$ | $HfYbErDyBaCuO_{10}$ |
| $YErGdLaTe_{10}$ | $YErLaDyNdTe_{10}$ | $NbYHfErGdBaO_{10}$ | $YbErSmLaDyGdTe_{10}$ |
| $YErLaNdTe_{10}$ | $YErDyGdLaTe_{10}$ | $NbYHfErBaScO_{10}$ | $YbErSmLaDyNdTe_{10}$ |
| $YSmLaDyTe_{10}$ | $YErDyLaNdTe_{10}$ | $NbYHfDyEuScO_{10}$ | $YbErSmDyGdLaTe_{10}$ |
| $YSmLaGdTe_{10}$ | $YSmLaDyGdTe_{10}$ | $NbYHfDyGdBaO_{10}$ | $YbErSmDyLaNdTe_{10}$ |
| $YSmDyLaTe_{10}$ | $YSmLaDyNdTe_{10}$ | $NbYHfEuGdScO_{10}$ | |
| $YSmEuNdTe_{10}$ | $YSmDyEuGdTe_{10}$ | $NbHfErSmBaTiO_{10}$ | |
| $YSmGdLaTe_{10}$ | $YSmDyGdLaTe_{10}$ | $NbHfErDyBaTiO_{10}$ | |
| $YLaDyGdTe_{10}$ | $YSmDyLaNdTe_{10}$ | $NbHfErDyBaScO_{10}$ | |
| $YLaDyNdTe_{10}$ | $YbErSmLaDyTe_{10}$ | $NbHfErGdBaScO_{10}$ | |
| $YDyGdLaTe_{10}$ | $YbErSmLaGdTe_{10}$ | $NbHfErBaCeScO_{10}$ | |
| $YDyLaNdTe_{10}$ | $YbErSmLaNdTe_{10}$ | $NbHfSmDyBaTiO_{10}$ | |
| $YEuGdBaTe_{10}$ | $YbErSmDyLaTe_{10}$ | $NbHfSmEuBaTiO_{10}$ | |
| $HfTiScCuO_{10}$ | $YbErSmDyNdTe_{10}$ | $NbHfLaDyEuScO_{10}$ | |

*Table 4 XGBoost predicted HECs with RT-ultralow κ.*

| Compounds | κ W/(m·K) |
|---|---|
| $HfBaTe_{10}$ | 0.037 |
| $ZrNbHfO_{10}$ | 0.012 |
| $ZrDyGdTe_{10}$ | 0.059 |
| $NbHfLaTe_{10}$ | 0.032 |
| $NbHfLaTe_{10}$ | 0.032 |
| $NbGdCuO_{10}$ | 0.059 |
| $YSmDyTe_{10}$ | 0.032 |
| $SmDyCuTe_{10}$ | 0.045 |
| $YbSmDyGdTe_{10}$ | 0.061 |

*Table 5 XGBoost predicted HECs with RT-extremely low κ.*

### 3.2 Feature importance analysis

To gain a better understanding of the factors influencing the thermal conductivity of ceramics and HECs, feature analysis was conducted on both the Random Forest (RF) and XGBoost models. The importance of each feature was determined and ranked based on their contributions to the prediction performance. For RF, the importance scores were calculated by measuring the average impurity reduction associated with each feature across all decision trees in the random forest. In the case of XGBoost, the feature importance values represented the average gain of each feature across all trees in the boosting process.

The results of the feature analysis, as shown in Figure 19 and Figure 20, indicated that certain features played a more significant role in determining κ. Specifically, the Maximum Mendeleev Number, Minimum Total Unfilled States, and Number of Components emerged as the most important features among the 51 considered. Furthermore, it was observed that size disorder, which is a measure of the variation in ionic radii within a composition, played a crucial role in the RF model. This may ascribe to its important effects on κ of HECs, which we will explain later. To further investigate the relationship between these important features and the predicted thermal conductivity, we conducted analysis on the 24,179 HECs in the dataset.

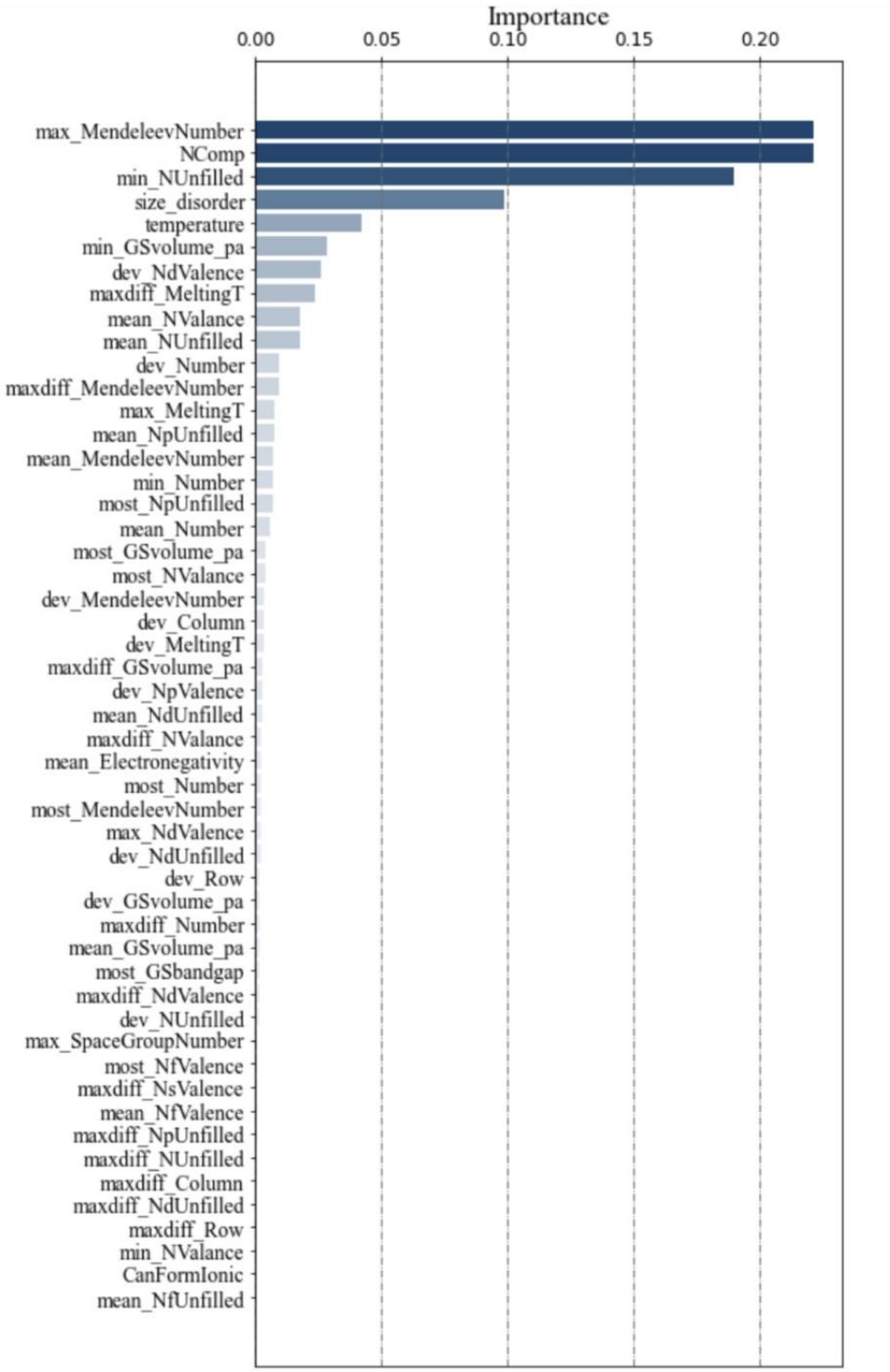


*Figure 19 Feature importance ranking of RF.*

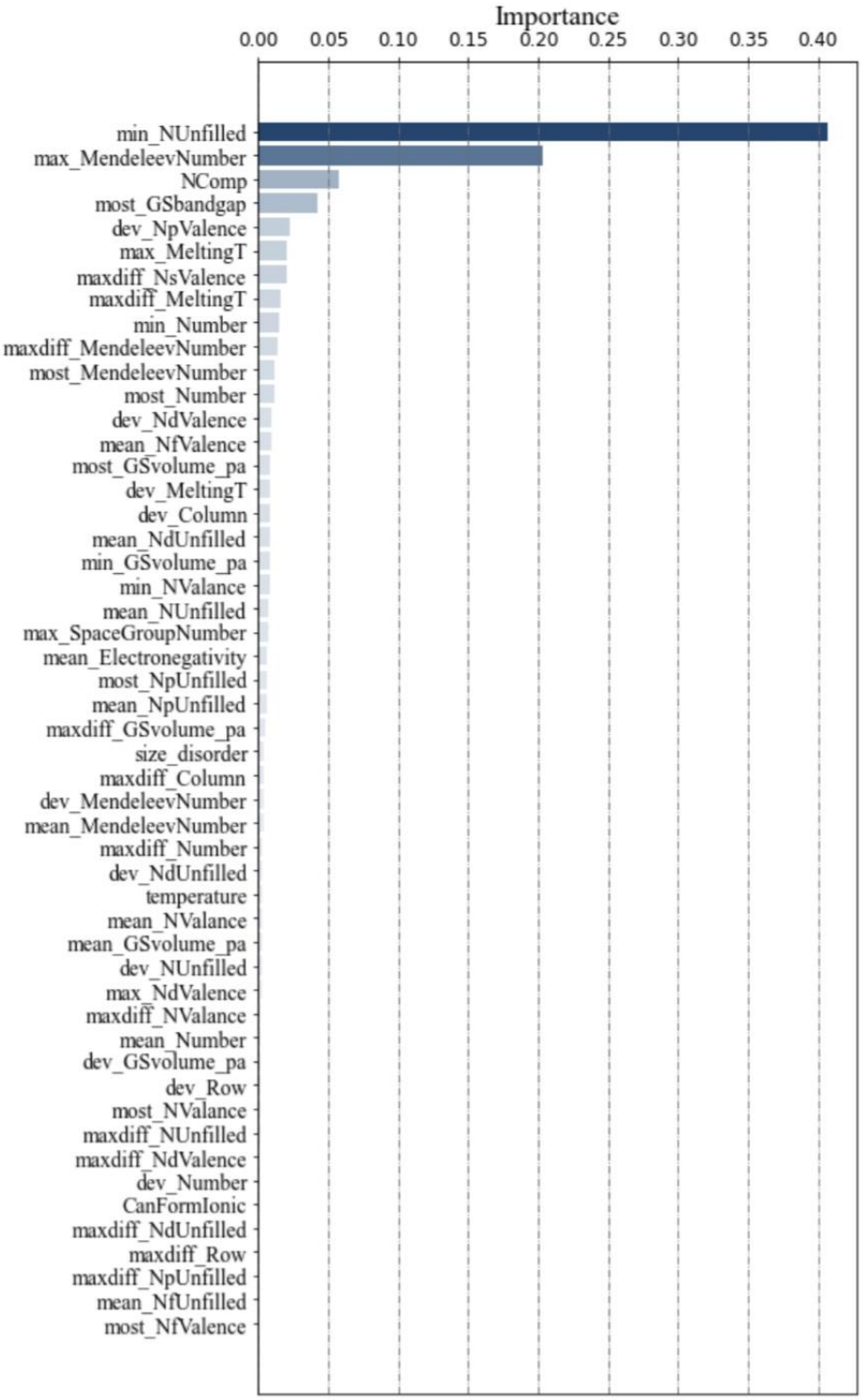


*Figure 20 Feature importance ranking of XGBoost.*

### 3.2.1 Maximum Mendeleev Number

In Figure 21, a clear trend is observed where an increase in the Maximum Mendeleev Number is associated with a decrease in the average thermal conductivity. The Maximum Mendeleev Number represents the anions in the compounds, specifically B, C, O, and Te, with corresponding values of 71, 77, 81, and 90, respectively. This trend can be explained by considering the effect of atomic size and mass on heat transport. As the Mendeleev number increases, it indicates larger atomic sizes and masses of the anions. This, in turn, leads to a reduction in the phonon mean free path of the heavier elements during the process of heat transport. The decrease in phonon mean free path restricts the movement of heat-carrying particles (phonons), resulting in a lower thermal conductivity. The results obtained in this study are consistent with previous research conducted on inorganic compounds. Wang et al. found that compounds containing light elements such as B, C, N, and O tend to exhibit higher thermal conductivity (κ) (Wang et al., 2020).

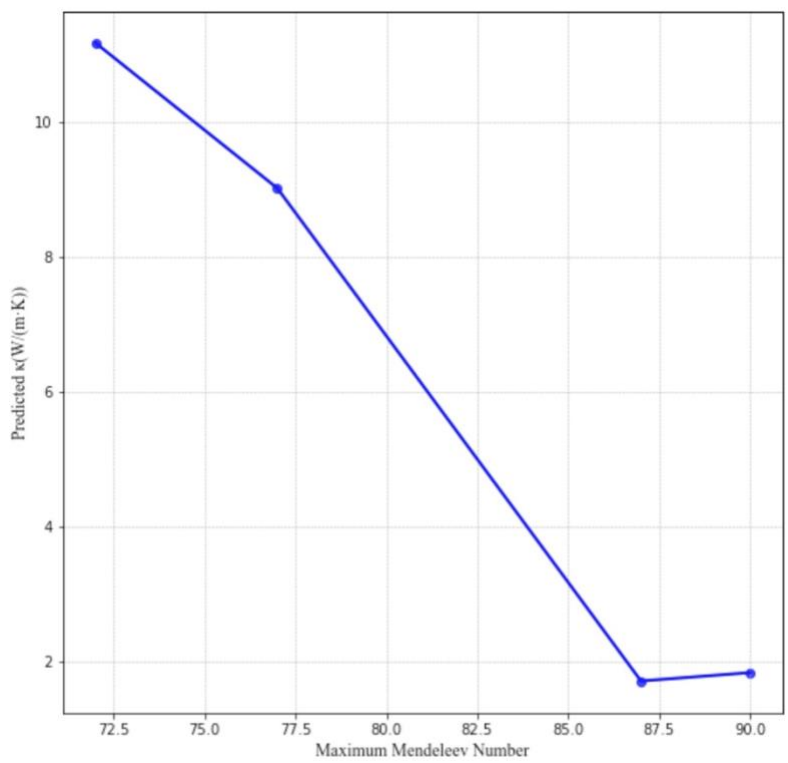


*Figure 21 Plot of XGBoost predicted κ (averaged) as function to Maximum Mendeleev Number.*

### 3.2.2 Minimum Total Unfilled States

Next, in Figure 22, we explore the relationship between the minimum total unfilled states and the average thermal conductivity. The results show a positive correlation between these two variables. This could be explained by Wiedemann-Franz Law $\kappa_e/\sigma = L_0T$, where $L_0$ is the Lorenz number and T is the temperature, $\kappa_e$ is the electron thermal conductivity and $\sigma$ is electronic conductivity. When there are more unfilled electronic states, there is a greater likelihood of finding mobile charge carriers, leading to enhanced electronic conductivity. Consequently, materials with higher electronic conductivity tend to exhibit higher thermal conductivity.

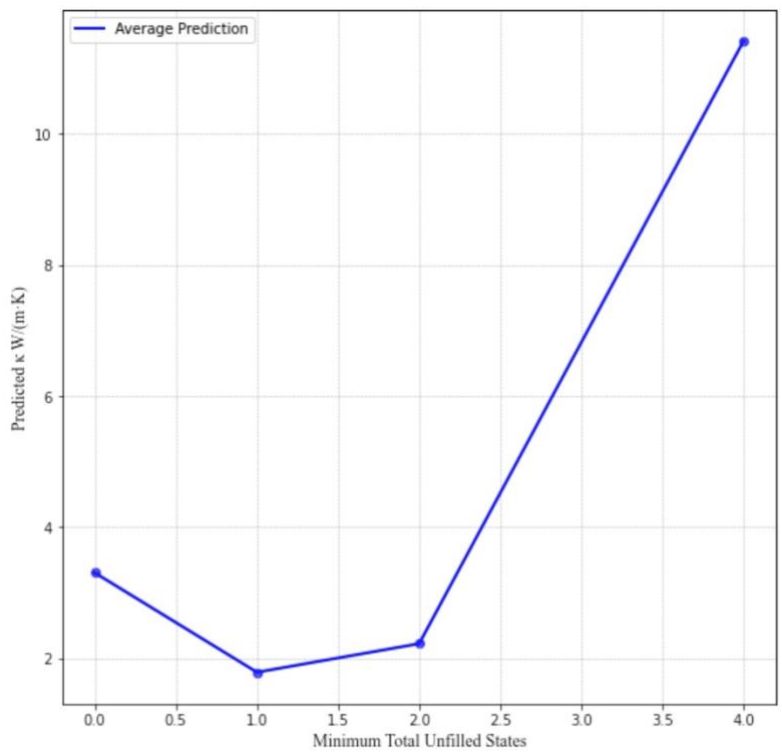


*Figure 22 Plot of XGBoost predicted κ (averaged) as function to Minimum Total Unfilled States.*

### 3.2.3 Number of Components

Additionally, our findings indicate that an increasing number of components in ceramics generally leads to a decrease in conductivity, as depicted in Figure 23. Notably, there is a significant drop when the number of components increases to four, which suggests the reduced κ in multicomponent ceramics. Liu et al. established a relationship between configurational entropy and lattice thermal conductivity by analyzing multiple multicomponent compounds (Liu et al., 2017). They attributed the reduction in lattice thermal conductivity to significant mass and strain fluctuations occurring among the different components. Consequently, an increasing number of components in ceramics result in higher configurational entropy and lower lattice thermal conductivity due to the increased complexity and disorder introduced by the additional components.

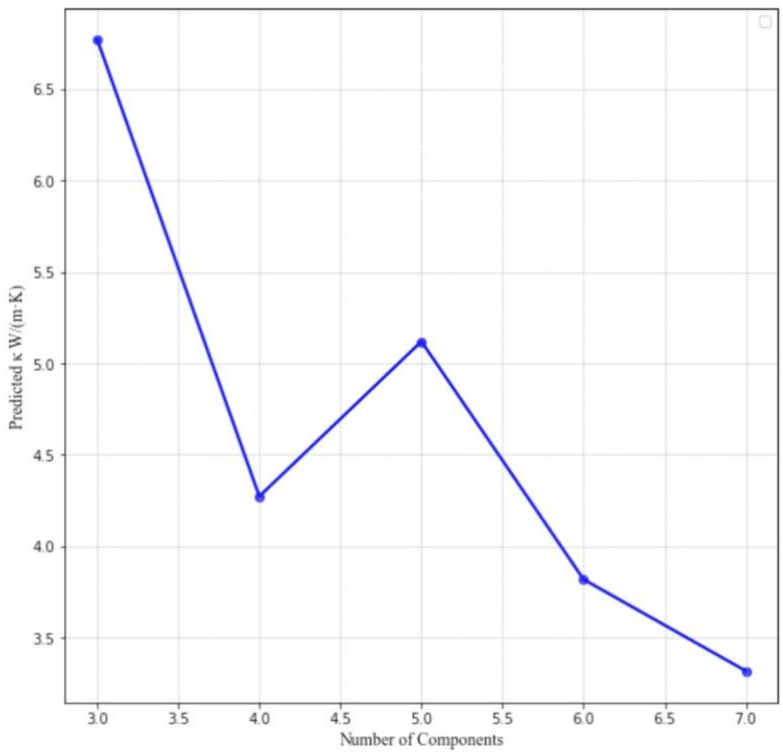


*Figure 23 Plot of XGBoost predicted κ (averaged) as function to Number of Components.*

### 3.2.4 Size Disorder in Lowering κ of HECs

Size disorder can lead to distorted sublattices and oversized cages, which in turn can strongly suppress heat conduction (Braun et al., 2018; Yang et al., 2019). Previous studies have also reported size disorder described by the δ parameter is a key descriptor controlling the reduced thermal conductivity in HECs (Andrew et al., 2020). To show that, the thermal conductivity of RT-ultralow HECs predicted by the ML model in this work were represented by the δ parameter as shown in Figure 24. A general trend, in which the κ increases by increasing the size disorder as measured by the δ parameter which further validates the capability of our model in prediction κ of HECs.

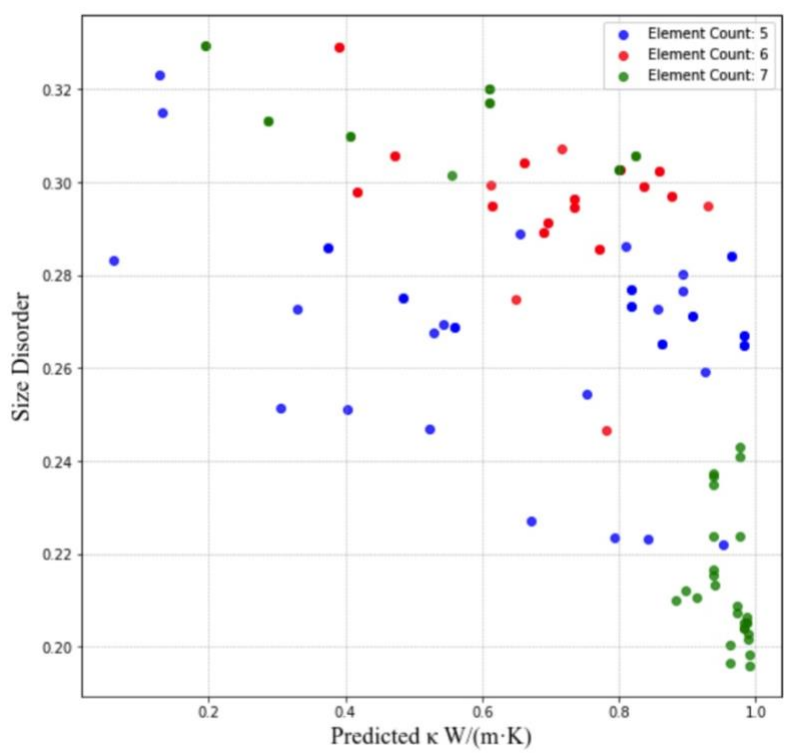


*Figure 24 Scatter of XGBoost predicted κ (averaged) with Size Disorder on RT-ultralow HECs.*

### 3.3 Limitations

One limitation of our study is the absence of structural features such as space group, lattice type, and the number of species. These features can provide valuable insights into the temperature dependence behavior of thermal conductivity in materials. Wang et al. (2020) demonstrated the significance of incorporating crystal properties and composition-weighted properties as descriptors in machine learning models for accurate thermal conductivity predictions using data from the Inorganic Crystallographic Structure Database. By considering these structural features, which capture important aspects of atomic arrangement and bonding, a more comprehensive understanding of the thermal transport mechanisms can be achieved. In future investigations, it will be beneficial to develop descriptors that effectively represent the structural features of high-entropy ceramics, thus enabling a more robust analysis of the temperature-dependent behavior of κ.

Overall, by analyzing the feature importance, we can gain insight into the descriptors that affect κ of ceramics, especially HECs. In addition, all these findings will provide further validation on the built ML model, where its prediction outcomes are fairly consistent with materials science concepts.

# CHAPTER 4

# CONCLUSION

In this study, we employed machine learning (ML) models to explore novel high-entropy ceramics (HECs) with low thermal conductivity. Our input features comprised compositional attributes calculated based on elemental properties, stoichiometry, valence orbital occupation, and ionic compound characteristics. Additionally, we included size disorder as a descriptor due to its impact on the thermal conductivity (κ) of HECs. The Random Forest, XGBoost, and Kernel Ridge Regression models exhibited robust performance and accurately reproduced the temperature-κ behavior observed in experimental data for HECs. Notably, the KRR model effectively captured the influence of anion content on thermal conductivity.

Through feature importance analysis, we determined that the Maximum Mendeleev Number, Minimum Total Unfilled States, and Number of Components were the most significant features influencing the thermal conductivity of HECs. Leveraging the XGBoost model, we explored a wide composition space of HEC-based materials, enabling the identification of promising compositions with desirable thermal conductivity properties. To validate our model further, we examined the relationship between the important features and κ values of the generated HECs. Our results aligned with previous findings, particularly highlighting the strong correlation between size disorder and reduced κ in HECs.

Overall, our study validates the efficacy of ML in the exploration and design of HECs with low thermal conductivity. Additionally, it provides valuable insights into the mechanisms underlying the unique thermal properties of HECs.

## REFERENCES


Albedwawi, S. H., AlJaberi, A., Haidemenopoulos, G. N., & Polychronopoulou, K. (2021). High entropy oxides-exploring a paradigm of promising catalysts: A review. Materials & Design, 202(109534), 109534. https://doi.org/10.1016/j.matdes.2021.109534

Arbuzova, T. I., Naumov, S. V., Kozlov, E. A., Arbuzov, V. L., Shal'nov, K. V., Gizhevskiĭ, B. A., Chukalkin, Y. G., & Voronin, V. I. (2006). Influence of structural defects on magnetic properties of the submicrometer ceramic $Mn_3O_4$. Journal of Experimental and Theoretical Physics, 102(6), 931–937. https://doi.org/10.1134/s1063776106060069

Braun, J. L., Rost, C. M., Lim, M., Giri, A., Olson, D. H., Kotsonis, G. N., Stan, G., Brenner, D. W., Maria, J.-P., & Hopkins, P. E. (2018). Charge-induced disorder controls the thermal conductivity of entropy-stabilized oxides. Advanced Materials (Deerfield Beach, Fla.), 30(51), e1805004. https://doi.org/10.1002/adma.201805004

Carrete, J., Li, W., Mingo, N., Wang, S., & Curtarolo, S. (2014). Finding unprecedentedly low-thermal-conductivity half-Heusler semiconductors via high-throughput materials modeling. Physical Review. X, 4(1). https://doi.org/10.1103/physrevx.4.011019

Chen, Lei, Zhang, W., Lu, W., Wei, B., Huo, S., Wang, Y., & Zhou, Y. (2022). Low thermal conductivity of dense $(TiZrHfVNbTa)C_x$ high-entropy carbides by tailoring carbon stoichiometry. Journal of Advanced Ceramics. https://doi.org/10.26599/jac.2023.9220665

Chen, Lihua, Tran, H., Batra, R., Kim, C., & Ramprasad, R. (2019). Machine learning models for the lattice thermal conductivity prediction of inorganic materials. Computational Materials Science, 170(109155), 109155. https://doi.org/10.1016/j.commatsci.2019.109155

Chen, Z., Tian, Z., Zheng, L., Ming, K., Ren, X., Wang, J., & Li, B. (2022). $(Ho_{0.25}Lu_{0.25}Yb_{0.25}Eu_{0.25})_2SiO_5$ high-entropy ceramic with low thermal conductivity, tunable thermal expansion coefficient, and excellent resistance to CMAS corrosion. Journal of Advanced Ceramics, 11(8), 1279–1293. https://doi.org/10.1007/s40145-022-0609-z

Clarke, D. R., & Phillpot, S. R. (2005). Thermal barrier coating materials. Materials Today (Kidlington,

England), 8(6), 22–29. https://doi.org/10.1016/s1369-7021(05)70934-2

Gild, J., Zhang, Y., Harrington, T., Jiang, S., Hu, T., Quinn, M. C., Mellor, W. M., Zhou, N., Vecchio, K., & Luo, J. (2016). High-entropy metal diborides: A new class of high-entropy materials and a new type of ultrahigh temperature ceramics. Scientific Reports, 6(1). https://doi.org/10.1038/srep37946

Giri, A., Braun, J. L., & Hopkins, P. E. (2018). Reduced dependence of thermal conductivity on temperature and pressure of multi-atom component crystalline solid solutions. Journal of Applied Physics, 123(1), 015106. https://doi.org/10.1063/1.5010337

Lim, M., Rak, Z., Braun, J. L., Rost, C. M., Kotsonis, G. N., Hopkins, P. E., Maria, J.-P., & Brenner, D. W. (2019). Influence of mass and charge disorder on the phonon thermal conductivity of entropy stabilized oxides determined by molecular dynamics simulations. Journal of Applied Physics, 125(5), 055105. https://doi.org/10.1063/1.5080419

Liu, J., Shao, G., Liu, D., Chen, K., Wang, K., Ma, B., Ren, K., & Wang, Y. (2020). Design and synthesis of chemically complex ceramics from the perspective of entropy. Materials Today. Advances, 8(100114), 100114. https://doi.org/10.1016/j.mtadv.2020.100114

Liu, R., Chen, H., Zhao, K., Qin, Y., Jiang, B., Zhang, T., Sha, G., Shi, X., Uher, C., Zhang, W., & Chen, L. (2017). Entropy as a gene-like performance indicator promoting thermoelectric materials. Advanced Materials (Deerfield Beach, Fla.), 29(38), 1702712. https://doi.org/10.1002/adma.201702712

Meredig, B., Agrawal, A., Kirklin, S., Saal, J. E., Doak, J. W., Thompson, A., Zhang, K., Choudhary, A., & Wolverton, C. (2014). Combinatorial screening for new materials in unconstrained composition space with machine learning. Physical Review. B, Condensed Matter and Materials Physics, 89(9). https://doi.org/10.1103/physrevb.89.094104

Moskovskikh, D., Vorotilo, S., Buinevich, V., Sedegov, A., Kuskov, K., Khort, A., Shuck, C., Zhukovskyi, M., & Mukasyan, A. (2020). Extremely hard and tough high entropy nitride ceramics. Scientific Reports, 10(1), 19874. https://doi.org/10.1038/s41598-020-76945-y

Oses, C., Toher, C., & Curtarolo, S. (2020). High-entropy ceramics. Nature Reviews. Materials, 5(4), 295–

309. https://doi.org/10.1038/s41578-019-0170-8

Padture, N. P., Gell, M., & Jordan, E. H. (2002). Thermal barrier coatings for gas-turbine engine applications. Science (New York, N.Y.), 296(5566), 280–284. https://doi.org/10.1126/science.1068609

Sang, W., Zhang, H., Zhang, H., Liu, X., Chen, X., Xie, W., Hou, R., Ma, H., Li, S., Wang, N., & Li, X. (2023). Novel $(Sm_{0.2}Lu_{0.2}Yb_{0.2}Y_{0.2}Dy_{0.2})_3TaO_7$ high-entropy ceramic for thermal barrier coatings. Ceramics International, 49(6), 9052–9059. https://doi.org/10.1016/j.ceramint.2022.11.061

Seko, A., Togo, A., Hayashi, H., Tsuda, K., Chaput, L., & Tanaka, I. (2015). Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization. Physical Review Letters, 115(20), 205901. https://doi.org/10.1103/PhysRevLett.115.205901

Shannon, R. D. (1976). Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. Acta Crystallographica. Section A, Crystal Physics, Diffraction, Theoretical and General Crystallography, 32(5), 751–767. https://doi.org/10.1107/s0567739476001551

Ward, L., Agrawal, A., Choudhary, A., & Wolverton, C. (2016). A general-purpose machine learning framework for predicting properties of inorganic materials. Npj Computational Materials, 2(1). https://doi.org/10.1038/npjcompumats.2016.28

Wright, A. J., Wang, Q., Ko, S.-T., Chung, K. M., Chen, R., & Luo, J. (2020). Size disorder as a descriptor for predicting reduced thermal conductivity in medium- and high-entropy pyrochlore oxides. Scripta Materialia, 181, 76–81. https://doi.org/10.1016/j.scriptamat.2020.02.011

Yang, J., Qian, X., Pan, W., Yang, R., Li, Z., Han, Y., Zhao, M., Huang, M., & Wan, C. (2019). Diffused lattice vibration and ultralow thermal conductivity in the binary Ln-Nb-O oxide system. Advanced Materials (Deerfield Beach, Fla.), 31(24), e1808222. https://doi.org/10.1002/adma.201808222

Zhang, R.-Z., & Reece, M. J. (2019). Review of high entropy ceramics: design, synthesis, structure and properties. Journal of Materials Chemistry. A, Materials for Energy and Sustainability, 7(39), 22148–

22162. https://doi.org/10.1039/c9ta05698j

Zhang, Y., Zhou, Y. J., Lin, J. P., Chen, G. L., & Liaw, P. K. (2008). Solid-solution phase formation rules for multi-component alloys. Advanced Engineering Materials, 10(6), 534–538. https://doi.org/10.1002/adem.200700240

## REFERENCES – DATASET


Aggarwal, R. L., Ripin, D. J., Ochoa, J. R., & Fan, T. Y. (2005). Thermo-optic properties of laser crystals in the 100-300 K temperature range: $Y_3Al_5O_{12}$ (YAG), $YAlO_3$ (YALO) and $LiYF_4$ (YLF). In H. J. Hoffman & R. K. Shori (Eds.), Solid State Lasers XIV: Technology and Devices. SPIE.

Braun, J. L., Rost, C. M., Lim, M., Giri, A., Olson, D. H., Kotsonis, G. N., Stan, G., Brenner, D. W., Maria, J.-P., & Hopkins, P. E. (2018). Charge-induced disorder controls the thermal conductivity of entropy-stabilized oxides. Advanced Materials (Deerfield Beach, Fla.), 30(51), 1805004. https://doi.org/10.1002/adma.201805004

Chen, H., Zhao, Z., Xiang, H., Dai, F.-Z., Xu, W., Sun, K., Liu, J., & Zhou, Y. (2020). High entropy $(Y_{0.2}Yb_{0.2}Lu_{0.2}Eu_{0.2}Er_{0.2})_3Al_5O_{12}$: A novel high temperature stable thermal barrier material. Journal of Materials Science & Technology, 48, 57–62. https://doi.org/10.1016/j.jmst.2020.01.056

Chen, K., Pei, X., Tang, L., Cheng, H., Li, Z., Li, C., Zhang, X., & An, L. (2018). A five-component entropy-stabilized fluorite oxide. Journal of the European Ceramic Society, 38(11), 4161–4164. https://doi.org/10.1016/j.jeurceramsoc.2018.04.063

Du, A., Wan, C., Qu, Z., & Pan, W. (2009). Thermal conductivity of monazite-type $REPO_4$(RE=la, Ce, Nd, Sm, Eu, Gd). Journal of the American Ceramic Society. American Ceramic Society, 92(11), 2687–2692. https://doi.org/10.1111/j.1551-2916.2009.03244.x

Gao, J., Zhao, L., Liu, Q., Wang, X., Zhang, S., & Li, J.-F. (2018). Antiferroelectric-ferroelectric phase transition in lead-free $AgNbO_3$ ceramics for energy storage applications. Journal of the American Ceramic Society. American Ceramic Society, 101(12), 5443–5450. https://doi.org/10.1111/jace.15780

Gild, J., Braun, J., Kaufmann, K., Marin, E., Harrington, T., Hopkins, P., Vecchio, K., & Luo, J. (2019). A high-entropy silicide: $(Mo_{0.2}Nb_{0.2}Ta_{0.2}Ti_{0.2}W_{0.2})Si_2$. Journal of Materiomics, 5(3), 337–343. https://doi.org/10.1016/j.jmat.2019.03.002

Gild, J., Samiee, M., Braun, J. L., Harrington, T., Vega, H., Hopkins, P. E., Vecchio, K., & Luo, J. (2018). High-entropy fluorite oxides. Journal of the European Ceramic Society, 38(10), 3578–3584.

https://doi.org/10.1016/j.jeurceramsoc.2018.04.010

Gild, J., Wright, A., Quiambao-Tomko, K., Qin, M., Tomko, J. A., Shafkat bin Hoque, M., Braun, J. L., Bloomfield, B., Martinez, D., Harrington, T., Vecchio, K., Hopkins, P. E., & Luo, J. (2020). Thermal conductivity and hardness of three single-phase high-entropy metal diborides fabricated by borocarbothermal reduction and spark plasma sintering. Ceramics International, 46(5), 6906–6913. https://doi.org/10.1016/j.ceramint.2019.11.186

Hu, L., Zhang, Y., Wu, H., Li, J., Li, Y., Mckenna, M., He, J., Liu, F., Pennycook, S. J., & Zeng, X. (2018). Entropy engineering of SnTe: Multi-principal-element alloying leading to ultralow lattice thermal conductivity and state-of-the-art thermoelectric performance. Advanced Energy Materials, 8(29), 1802116. https://doi.org/10.1002/aenm.201802116

Li, R.-Y., & Duan, Y.-H. (2016). Anisotropic elastic properties of MB (M = Cr, Mo, W) monoborides: a first-principles investigation. Philosophical Magazine (Abingdon, England), 96(10), 972–990. https://doi.org/10.1080/14786435.2016.1149247

Liu, D., Shi, B., Geng, L., Wang, Y., Xu, B., & Chen, Y. (2022). High-entropy rare-earth zirconate ceramics with low thermal conductivity for advanced thermal-barrier coatings. Journal of Advanced Ceramics, 11(6), 961–973. https://doi.org/10.1007/s40145-022-0589-z

Liu, H. (2018). Origin of the intermediate phase in lead zirconate, $PbZrO_3$. Journal of the American Ceramic Society. American Ceramic Society, 101(11), 5281–5286. https://doi.org/10.1111/jace.15786

Liu, R., Chen, H., Zhao, K., Qin, Y., Jiang, B., Zhang, T., Sha, G., Shi, X., Uher, C., Zhang, W., & Chen, L. (2017). Entropy as a gene-like performance indicator promoting thermoelectric materials. Advanced Materials (Deerfield Beach, Fla.), 29(38), 1702712. https://doi.org/10.1002/adma.201702712

Liu, Y., Jia, D., Zhou, Y., Zhou, Y., Zhao, J., Nian, H., & Liu, B. (2020). $Zn_{0.1Ca0.1}Sr_{0.4}Ba_{0.4}ZrO_3$: A non-equimolar multicomponent perovskite ceramic with low thermal conductivity. Journal of the European Ceramic Society, 40(15), 6272–6277. https://doi.org/10.1016/j.jeurceramsoc.2020.07.054

Moskovskikh, D. O., Vorotilo, S., Sedegov, A. S., Kuskov, K. V., Bardasova, K. V., Kiryukhantsev-

korneev, P. V., Zhukovskyi, M., & Mukasyan, A. S. (2020). High-entropy (HfTaTiNbZr)C and (HfTaTiNbMo)C carbides fabricated through reactive high-energy ball milling and spark plasma sintering. Ceramics International, 46(11), 19008–19014. https://doi.org/10.1016/j.ceramint.2020.04.230

Neshpor, V. S. (1968). The thermal conductivity of the silicides of transition metals. Journal of Engineering Physics, 15(2), 750–752. https://doi.org/10.1007/bf00829703

Qiu, S., Li, M., Shao, G., Wang, H., Zhu, J., Liu, W., Fan, B., Xu, H., Lu, H., Zhou, Y., & Zhang, R. (2021). (Ca,Sr,Ba)$ZrO_3$: A promising entropy-stabilized ceramic for titanium alloys smelting. Journal of Materials Science & Technology, 65, 82–88. https://doi.org/10.1016/j.jmst.2020.05.033

Ren, K., Wang, Q., Shao, G., Zhao, X., & Wang, Y. (2020). Multicomponent high-entropy zirconates with comprehensive properties for advanced thermal barrier coating. Scripta Materialia, 178, 382–386. https://doi.org/10.1016/j.scriptamat.2019.12.006

Ren, X., Tian, Z., Zhang, J., & Wang, J. (2019). Equiatomic quaternary $(Y_{1/4}Ho_{1/4}Er_{1/4}Yb_{1/4})_2SiO_5$ silicate: A perspective multifunctional thermal and environmental barrier coating material. Scripta Materialia, 168, 47–50. https://doi.org/10.1016/j.scriptamat.2019.04.018

Ridley, M., Gaskins, J., Hopkins, P., & Opila, E. (2020). Tailoring thermal properties of multi-component rare earth monosilicates. Acta Materialia, 195, 698–707. https://doi.org/10.1016/j.actamat.2020.06.012

Sharma, Y., Musico, B. L., Gao, X., Hua, C., May, A. F., Herklotz, A., Rastogi, A., Mandrus, D., Yan, J., Lee, H. N., Chisholm, M. F., Keppens, V., & Ward, T. Z. (2018). Single-crystal high entropy perovskite oxide epitaxial films. Physical Review Materials, 2(6). https://doi.org/10.1103/physrevmaterials.2.060404

Shi, M., Zhu, C., Lu, M., Meng, X., & Wei, M. (2018). Dy and Eu activated $Ca_3B_2O_6$ phosphors for near ultraviolet-based light-emitting diodes. Journal of the American Ceramic Society. American Ceramic Society, 101(12), 5461–5468. https://doi.org/10.1111/jace.15782

Tye, R. P., & Clougherty, E. V. (1970). Thermal and electrical conductivities of some electrically

conducting compounds. Pp 396-401 of Proceedings of the Fifth Symposium on Thermophysical Properties. /Bonilla, C. F. (Ed.). New York American Society of Mechanical Engineers (1970). https://www.osti.gov/biblio/4744096

Vaßen, R., Jarligo, M. O., Steinke, T., Mack, D. E., & Stöver, D. (2010). Overview on advanced thermal barrier coatings. Surface & Coatings Technology, 205(4), 938–942. https://doi.org/10.1016/j.surfcoat.2010.08.151

Vasudévan, A. K., & Petrovic, J. J. (1992). A comparative overview of molybdenum disilicide composites. Materials Science & Engineering. A, Structural Materials: Properties, Microstructure and Processing, 155(1–2), 1–17. https://doi.org/10.1016/0921-5093(92)90308-n

Wang, C., Zhang, W., Han, Z., & Shi, L. (2018). Behaviors of helium in $Cr_2AlC$ from first principles. Journal of the American Ceramic Society. American Ceramic Society, 101(12), 5771–5780. https://doi.org/10.1111/jace.15781

Wang, X., Xiang, H., Sun, X., Liu, J., Hou, F., & Zhou, Y. (2014). Thermal properties of a prospective thermal barrier material: $Yb_3Al_5O_{12}$. Journal of Materials Research, 29(22), 2673–2681. https://doi.org/10.1557/jmr.2014.319

Xing, J., Foroughi, P., Franco Hernandez, A., Behrens, A., & Cheng, Z. (2018). Facile one-step high-temperature spray pyrolysis route toward metal carbide nanopowders. Journal of the American Ceramic Society. American Ceramic Society, 101(12), 5323–5334. https://doi.org/10.1111/jace.15785

Yan, X., Constantin, L., Lu, Y., Silvain, J.-F., Nastasi, M., & Cui, B. (2018). $(Hf_{0.2}Zr_{0.2}Ta_{0.2}Nb_{0.2}Ti_{0.2})C$ high-entropy ceramics with low thermal conductivity. Journal of the American Ceramic Society. American Ceramic Society, 101(10), 4486–4491. https://doi.org/10.1111/jace.15779

Ye, B., Wen, T., Nguyen, M. C., Hao, L., Wang, C.-Z., & Chu, Y. (2019). First-principles study, fabrication and characterization of $(Zr_{0.25}Nb_{0.25}Ti_{0.25}V_{0.25})C$ high-entropy ceramics. Acta Materialia, 170, 15–23. https://doi.org/10.1016/j.actamat.2019.03.021

Zhao, Z., Chen, H., Xiang, H., Dai, F.-Z., Wang, X., Peng, Z., & Zhou, Y. (2019).

$(La_{0.2}Ce_{0.2}Nd_{0.2}Sm_{0.2}Eu_{0.2})PO4$: A high-entropy rare-earth phosphate monazite ceramic with low thermal conductivity and good compatibility with $Al_2O_3$. Journal of Materials Science & Technology, 35(12), 2892–2896. https://doi.org/10.1016/j.jmst.2019.08.012

Zhao, Z., Chen, H., Xiang, H., Dai, F.-Z., Wang, X., Xu, W., Sun, K., Peng, Z., & Zhou, Y. (2020a). High-entropy $(Y_{0.2}Nd_{0.2}Sm_{0.2}Eu_{0.2}Er_{0.2})AlO_3$: A promising thermal/environmental barrier material for oxide/oxide composites. Journal of Materials Science & Technology, 47, 45–51. https://doi.org/10.1016/j.jmst.2020.02.011

Zhao, Z., Chen, H., Xiang, H., Dai, F.-Z., Wang, X., Xu, W., Sun, K., Peng, Z., & Zhou, Y. (2020b). $(Y_{0.25}Yb_{0.25}Er_{0.25}Lu_{0.25})_2(Zr_{0.5}Hf_{0.5})_2O_7$: A defective fluorite structured high entropy ceramic with low thermal conductivity and close thermal expansion coefficient to $Al_2O_3$. Journal of Materials Science & Technology, 39, 167–172. https://doi.org/10.1016/j.jmst.2019.08.018

Zhao, Z., Xiang, H., Dai, F.-Z., Peng, Z., & Zhou, Y. (2019a). $(La_{0.2}Ce_{0.2}Nd_{0.2}Sm_{0.2}Eu_{0.2})_2Zr_2O_7$: A novel high-entropy ceramic with low thermal conductivity and sluggish grain growth rate. Journal of Materials Science & Technology, 35(11), 2647–2651. https://doi.org/10.1016/j.jmst.2019.05.054

Zhao, Z., Xiang, H., Dai, F.-Z., Peng, Z., & Zhou, Y. (2019b). $(TiZrHf)P_2O_7$: An equimolar multicomponent or high entropy ceramic with good thermal stability and low thermal conductivity. Journal of Materials Science & Technology, 35(10), 2227–2231. https://doi.org/10.1016/j.jmst.2019.05.030

Zhu, J., Meng, X., Xu, J., Zhang, P., Lou, Z., Reece, M. J., & Gao, F. (2021). Ultra-low thermal conductivity and enhanced mechanical properties of high-entropy rare earth niobates ($RE_3NbO_7$, RE = Dy, Y, Ho, Er, Yb). Journal of the European Ceramic Society, 41(1), 1052–1057. https://doi.org/10.1016/j.jeurceramsoc.2020.08.070

Zimmermann, J. W., Hilmas, G. E., Fahrenholtz, W. G., Dinwiddie, R. B., Porter, W. D., & Wang, H. (2008). Thermophysical properties of $ZrB_2$and $ZrB_2$–SiC ceramics. Journal of the American Ceramic Society. American Ceramic Society, 91(5), 1405–1411. https://doi.org/10.1111/j.1551-2916.2008.02268.x