Design of AlCuNiPtMn high entropy alloy with desired properties assisted by Machine learning

Submitted in partial fulfilment of the requirements of the degree of
Bachelor of Technology by

G Yaswanth 21CHB0A17

Dekka Hymavathi 21CHB0A14

Namuduri Bhaskar 21CHB0B36

Jinugu Pavan 21CHB0A24

Supervisor Name: Dr. Phani Kumar

DEPARTMENT OF CHEMICAL ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY WARANGAL

MAY 2024

Approval Sheet

This Project Work entitled Design of AlCuNiPtMn high entropy alloy with desired properties assisted by Machine learning by G Yaswanth_21CHB0A17, Dekka Hymavathi_21CHB0A14, Namuduri Bhaskar_21CHB0B36, Jinugu Pavan_21CHB0A24 is approved for the degree of BTech in Chemical Engineering.

Examiners

_____

_____

_____

_____

_____

Supervisor

_____

Chairman

_____

Date: _____ Place: _____

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Roll No.	Name of the student	Signature of the student
21CHB0A17	G Yaswanth	
21CHB0A14	Dekka Hymavathi	
21CHB0B36	Namuduri Bhaskar	
21CHB0A24	Jinugu Pavan	

Date: 02-05-2024

Certificate

This is to certify that the Project work entitled **"Design of AlCuNiPtMn high entropy alloy with desired properties assisted by Machine learning"** is a Bonafide record of work carried out by **"G Yaswanth_21CHB0A17 , Dekka Hymavathi_21CHB0A14 , Namuduri Bhaskar_21CHB0B36 , Jinugu Pavan_21CHB0A24"**, submitted to the faculty of Department of Chemical Engineering, in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Chemical Engineering at National Institute of Technology, Warangal during the academic year 2023-24.

Dr. P V Suresh,
Head of the Department,
Department of Chemical Engineering,
NIT Warangal.

Dr. Phani Kumar,
Assistant Professor,
Department of Chemical Engineering,
NIT Warangal.

ACKNOWLEDGEMENT

First and foremost, we would like to thank our project guide Dr. Phani Kumar for giving valuable insights about previously done research related to our topic (Design of AlCuNiPtMn high entropy alloy with desired properties assisted by Machine learning) which we believe was an excellent introduction for us to get our desired data for the HEA (high entropy alloy) properties to design AlCuNiPtMn using Machine learning. His suggestions greatly improved the quality of the project. We would never be able to finish the project without his support. We would also like to show our gratitude to our fellow mates, who were constantly there to clarify our doubts and recommend improvements to the project. Their help and support were of immense value to us.

LIST OF CONTENTS

LIST OF TABLES

ABSTRACT

High Entropy Alloys (HEAs) represent a novel class of materials composed of multiple elements in roughly equiatomic proportions, exhibiting exceptional mechanical and physical properties. This project focuses on harnessing machine learning techniques to predict the bulk modulus of HEAs, which is crucial for understanding their mechanical behavior and facilitating their application in various engineering domains.

In this study, we employ a machine learning approach to predict the bulk modulus of HEAs comprising 4 to 6 elements. A classifier is first utilized to identify whether a given HEA composition exists within the dataset, leveraging the distinctive characteristics of HEAs. Subsequently, an Extreme Gradient Boosting (XGBoost) regression model is employed to predict the bulk modulus for identified HEA compositions.

The dataset utilized for model training and evaluation encompasses a diverse range of HEA compositions along with their corresponding bulk moduli, sourced from experimental measurements and theoretical calculations. Feature engineering techniques tailored to HEAs are applied to extract relevant features from the elemental compositions.The performance of the developed model is assessed using various metrics, including mean absolute error, mean squared error, and R-squared. Additionally, the model's robustness is evaluated through cross-validation and sensitivity analysis.

The proposed machine learning framework not only streamlines the prediction process but also provides insights into the complex relationships between alloy composition and bulk modulus in HEAs. This predictive model can serve as a valuable tool for materials scientists and engineers in the design and optimization of HEAs with tailored mechanical properties, thereby accelerating the development of advanced materials in engineering applications.

Chapter 1

Introduction

1.1 What is HEA ?

High-entropy alloys (HEAs) appeared at the beginning of the 21st century as a new class of materials. These are the alloys that are formed by mixing equal or relatively large proportions of (usually) five or more elements. The term "high-entropy alloys" was coined because the entropy increase of mixing is substantially higher when there is a larger number of elements in the mix, and their proportions are more nearly equal.

HEAs compositions can exhibit deviations from equiatomic or near-equiatomic proportions while still being classified as HEAs. This is because the key defining characteristic of HEAs is their high configurational entropy, which is a result of the random arrangement of multiple principal elements in the alloy. Even in cases where one element is present in significantly larger quantities than others, and another element is present in smaller quantities, the alloy can still exhibit high entropy and form a solid solution. As long as the alloy contains multiple principal elements and demonstrates a high degree of disorder in its atomic arrangement, it can be considered an HEA, highlighting the flexibility and diversity of this class of materials.

1.2 What is Machine Learning?

Machine learning is a subset of artificial intelligence (AI) that involves the development of algorithms and statistical models that enable computers to learn and make predictions or decisions based on data, without being explicitly programmed for specific tasks.

Machine learning offers a promising approach to simplify Density Functional Theory (DFT) calculations and other large-scale simulations in predicting properties of alloys. By training machine learning models on datasets generated from DFT calculations, these models can learn the complex relationships between the composition, structure, and properties of alloys. Once trained, these models can provide accurate predictions of material properties in a fraction of the time and computational cost required for traditional simulations.

In the context of predicting the bulk modulus of high entropy alloys (HEAs), machine learning techniques are utilized to analyze large datasets of HEA compositions and their corresponding bulk moduli, allowing the algorithm to identify patterns and relationships within the data. These patterns are then used to train predictive models that can accurately estimate the bulk modulus of new or unseen HEA compositions. Machine learning algorithms such as Extreme Gradient Boosting (XGBoost) are particularly effective for this task due to their ability to handle complex interactions among multiple input features and their high predictive performance. By leveraging machine learning in materials science research, scientists and engineers can expedite the process of material discovery and optimization, leading to advancements in various engineering applications.

1.3 Applications of HEAs

- Biomedicine(as coatings)
- Corrosion resistance
- Refrigeration applications
- High temperature applications
- In power industry
- Catalysis
- Energy storage

Out of all these applications, HEAs application in catalysis grabbed more attention. HEA's demonstrate remarkable performance in catalyzing reactions like oxygen reduction reaction , ammonia decomposition and more, offering versatile solutions for clean energy conversion.The integration of HEA's in electro/thermo-catalytic systems holds promise for enhancing efficiency and reducing environmental impact in energy conversion processes. As HEA's continue to evolve, their role in catalytic processes is poised to expand, driving innovations towards sustainable and efficient energy conversion technologies.

Chapter 2

Literature Review

Introduction:

The literature we have gone through delves into the recent advancements in HEAs, focusing on their synthesis methodologies, catalytic applications, and future prospects. In their paper, "Recent progress in high-entropy alloys for catalysts: synthesis, applications, and prospects," K. Li and W. Chen provide a comprehensive review of the current state of research on high-entropy alloys as catalysts, highlighting their synthesis methods, applications, and future prospects.

2.1 Fundamentals of HEAs:

The authors began by elucidating the fundamental concepts of HEAs, highlighting their thermodynamic properties and the factors influencing phase selection. HEAs, characterized by their near-equiatomic composition of five or more elements, offer intriguing possibilities for enhancing material properties and performance**.**

2.2 Synthesis Strategies:

Various synthesis strategies for HEAs are explored, including carbothermal shock, fast-moving bed pyrolysis, electro-shock, and theoretical prediction methods. These innovative approaches pave the way for fabricating HEAs with tailored structures and compositions, enabling targeted applications in catalysis.

2.3 Applications:

The authors emphasized the catalytic potential of HEAs in clean energy conversion reactions such as hydrogen evolution reaction (HER), oxygen evolution reaction (OER), oxygen reduction reaction (ORR), carbon dioxide reduction reaction (CO2RR), and ammonia (NH3) decomposition. HEAs exhibit remarkable electrochemical performance, attributed to synergistic effects among multiple metal components, which provide diverse active sites for catalytic reactions.

2.4 Prospects and Future Opportunities:

Looking ahead, they went through the challenges and opportunities in the synthesis and application of HEAs. Computational methods, including density functional theory (DFT)

calculations and machine learning, are highlighted as valuable tools for designing novel HEAs with enhanced catalytic properties. The authors envision continued advancements in HEA research, driving innovation in catalytic processes for sustainable energy conversion.

2.5 Conclusion:

In summary, the papers provided a comprehensive overview of recent progress in the synthesis, applications, and prospects of HEAs for catalytic purposes. By elucidating fundamental principles, exploring synthesis strategies, and highlighting catalytic applications, the review underscores the pivotal role of HEAs in advancing clean energy technologies. As research in this field continues to evolve, HEAs hold promise for addressing the pressing challenges of energy conversion and environmental sustainability.

Chapter 3

Research Objective

HEAs useful properties gave them a wide range of applications in many industries. HEA application in catalysis is one of those. The catalytic activity can be enhanced by following an ideal composition and structure for formation of the HEA. The HEA of our interest is AlCuNiPtMn which is a nanoporous Pt based high entropy alloy (HEA) having an important application as catalyst in oxidation-reduction reactions (ORR) and in CO2 reduction to useful products**. Our main objective is Leveraging Machine Learning to Determine the ideal composition of the AlCuNiPtMn high entropy alloy (HEA) that exhibits the best bulk modulus for catalyst applications.**

Chapter 4

Methodology

4.1 Dataset Description

The dataset used in this study consists of two main components: the classifier dataset and the regressor dataset. The classifier dataset contains compositions of High Entropy Alloys (HEAs) with 14 unique elements in equiatomic proportions, along with labels indicating whether they form a valid HEA. This dataset has a shape of (1320, 14). The regression dataset includes compositions of HEAs with the same 14 unique elements, along with additional inputs such as crystal structure and lattice constant. The regressor dataset has a shape of (1146, 17).

4.2 Data Preprocessing

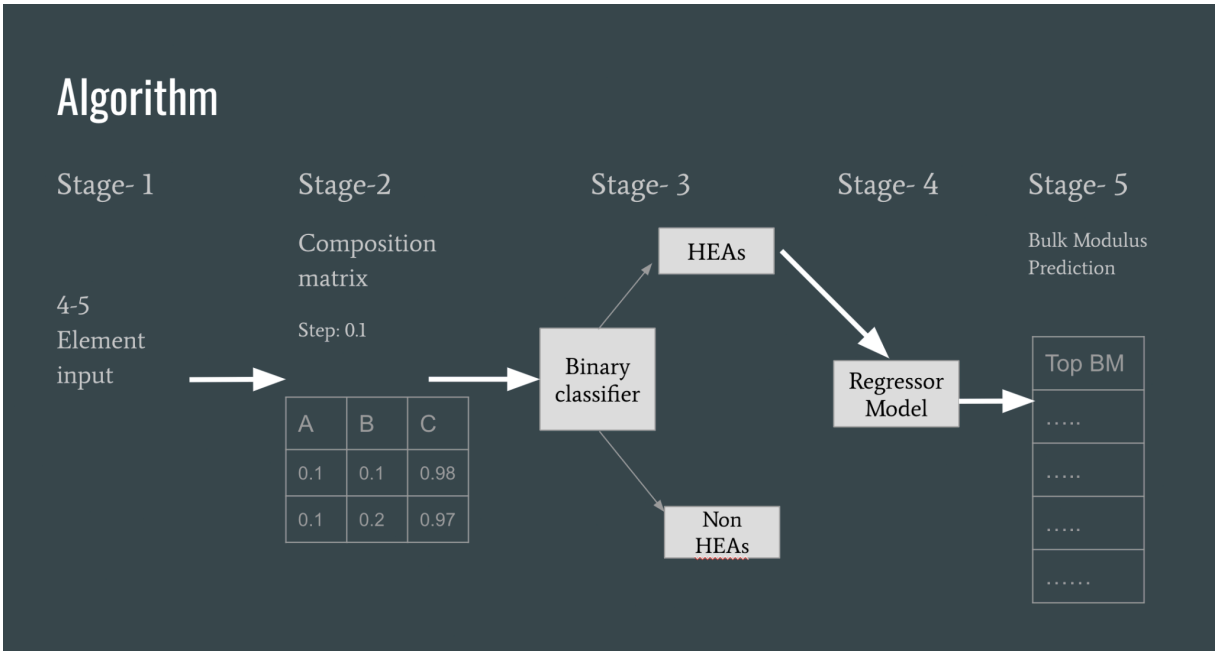
In the data preprocessing stage, relevant features are extracted from the elemental compositions using domain-specific knowledge about HEAs. Standard scaling is applied to normalize the features, ensuring that each feature contributes equally to the model training process. The datasets are then split into training and test sets using an 80-20 split to evaluate the model performance.

4.3 Classifier Model

The XGBoost classifier is selected as the best-performing model for identifying HEA compositions. It is trained on the classifier dataset to classify whether a given composition forms a valid HEA. The model is evaluated using metrics such as accuracy, precision, recall, and F1-score to assess its performance in classifying HEA compositions.

4.4 Regressor Model

For predicting the bulk modulus of HEAs, the Random Forest regressor is chosen as the model of choice. It is trained on the regressor dataset, which includes compositions of HEAs along with crystal structure and lattice constant as inputs. The model's performance is evaluated using metrics such as mean absolute error, mean squared error, and R-squared to determine its accuracy in predicting the bulk modulus of HEAs.



4.5 Predicting Best Composition

The classification model is employed to classify all possible combinations of given elements, determining whether they can form a High Entropy Alloy (HEA). Once the classifier identifies valid HEA compositions, these compositions are used as inputs for predicting the bulk modulus. A regression model, specifically the Random Forest regressor, is utilized for this task, trained on a dataset containing compositions of HEAs along with crystal structure and lattice constant as inputs. The regression model predicts the bulk modulus for each valid HEA composition, and the composition with the highest predicted bulk modulus is considered the final result.The algorithm of the same is pictured above. This integrated approach not only streamlines the prediction process but also provides insights into the mechanical behavior of HEAs, aiding in their application across various engineering domains.

Chapter 5

Results and Discussions

5.1 HEA classifier model results

The performance of various classifier models in classifying whether a certain composition of given elements can form a High Entropy Alloy (HEA) was evaluated using both train and test accuracies. The following table summarizes the results:

Model	Train Accuracy	Test Accuracy
Logistic Regression	0.6666666667	0.6684303351
Support Vector Classifier	0.8	0.7019400353
KNeighborsClassifier	0.7553030303	0.7231040564
Gradient Boosting Classifier	0.8340909091	0.7107583774
Random Forest Classifier	0.8113636364	0.7125220459
XGB Classifier	0.8348484848	0.7425044092
Decision Tree Classifier	0.7439393939	0.6613756614

The models were trained and tested using a dataset that included the composition of elements and their corresponding classification as HEA or non-HEA. The results showed varying levels of accuracy across different models. The Logistic Regression model achieved a train

accuracy of 0.667 and a test accuracy of 0.668, indicating a balanced performance. The Support Vector Classifier performed well on the training set with an accuracy of 0.800 but showed a slightly lower test accuracy of 0.702, suggesting some overfitting. In contrast, the KNeighborsClassifier demonstrated good performance with a higher test accuracy of 0.723, making it a suitable choice for this classification task. The Gradient Boosting Classifier showed high training accuracy of 0.834 but a lower test accuracy of 0.711, indicating potential overfitting. The Random Forest Classifier achieved a reasonable balance between training and test accuracies, with values of 0.811 and 0.713 respectively. The XGBClassifier emerged as the most effective model, with the highest test accuracy of 0.743, highlighting its potential for accurate prediction of HEA formation.

5.2 HEA property prediction model results

Model	Train R-square	Test R-square	Train MSE	Test MSE	Train MAE	Test MAE
Random Forest	0.9810378925	0.8121330504	23.00011274	199.7973046	3.126534844	8.21356331
XG Boost Regressor	0.9937363712	0.8408792451	7.597476654	169.2256036	1.8851371	7.592617916
Linear Regression	0.8070963635	0.7698966572	233.9827143	244.7158895	10.42174722	9.948377614
Ridge Regression	0.8069927057	0.7705069967	234.1084462	244.0667908	10.41190781	9.945974339
KNN Regressor	0.8308611116	0.7607565526	205.1572326	254.4364299	10.62040403	10.75973604

The Random Forest model achieved a high R-square of 0.981 on the training set and 0.812 on the test set, indicating a strong ability to explain the variance in the data. However, it exhibited a relatively high test MSE of 199.797 and MAE of 8.214, suggesting some level of overfitting. The XGBoost Regressor outperformed the Random Forest model with a higher test R-square of 0.841 and lower test MSE of 169.226 and MAE of 7.593, despite having a

slightly lower training R-square of 0.994. The Linear Regression and Ridge Regression models showed similar performance, with R-square values around 0.807 and comparable MSE and MAE values. The KNN Regressor demonstrated the highest R-square among the models, with 0.831 on the training set but performed relatively poorly on the test set with an R-square of 0.761 and higher MSE and MAE values compared to other models. Overall, the XGBoost Regressor showed the best performance in predicting the bulk modulus of alloys, indicating its potential for accurate modeling of material properties.

5.3 Top 10 Compositions of AlCuNiPtMn HEA

	Al	Cu	Ni	Pt	Mn	BM_pred
1	0.16	0.02	0.79	0.01	0.02	162.4233579
2	0.16	0.02	0.78	0.02	0.02	162.4200689
3	0.16	0.01	0.8	0.01	0.02	162.4111406
4	0.16	0.01	0.79	0.02	0.02	162.4078515
5	0.15	0.02	0.8	0.01	0.02	162.4019653
6	0.15	0.02	0.79	0.02	0.02	162.3986762
7	0.15	0.01	0.81	0.01	0.02	162.3725928
8	0.15	0.01	0.8	0.02	0.02	162.3693037
9	0.16	0.02	0.77	0.03	0.02	162.1878261
10	0.16	0.01	0.78	0.03	0.02	162.1756088

The compositions show a combination of Al, Cu, Ni, Pt, and Mn, with varying proportions. Despite not adhering strictly to the equiatomic composition rule, these alloys are likely to

exhibit high configurational entropy due to the presence of multiple principal elements. This high entropy contributes to their potential for enhanced mechanical and physical properties, making them suitable for various engineering applications. The Composition 1 (Al: 0.16, Cu: 0.02, Ni: 0.79, Pt: 0.01, Mn: 0.02) has a bulk modulus prediction of 162.4233579 GPa highest among all the possible combinations of the desired elements as predicted by the ML model.

Chapter 6

Summary and Conclusions

The composition of Al, Cu, Ni, Pt, Mn in the ratios 0.16, 0.02, 0.79, 0.01, 0.02 respectively results in a high bulk modulus due to several reasons:

- Phase Stability: The specific combination of elements in this composition may result in the formation of stable phases or intermetallic compounds that contribute to enhanced mechanical properties, including bulk modulus.
- Atomic Structure: Copper (Cu) and Nickel (Ni) are adjacent to each other in the periodic table, with atomic numbers 29 and 28 respectively. They have similar atomic structures and are completely miscible in both liquid and solid states
- Crystal Structure: Cu-Ni alloys crystallize in a face-centered cubic lattice over the whole concentration range, which contributes to their high bulk modulus

This research shows that machine learning models like the GBC classifier and Random Forest Regressor can be very useful in finding high entropy alloys (HEAs) with the best bulk modulus values. These predicted values help users choose the HEA with the highest bulk modulus, which is important for strong applications. This makes it easier for researchers to create HEAs with high bulk modulus in experiments.Additionally, this approach could be expanded to explore how other properties of HEAs relate to their elemental composition. Using machine learning to design virtual HEAs reduces the need for traditional experimental methods, offering a more efficient alternative. As more data becomes available, the accuracy of these models can improve significantly.

Chapter 7

Future Scope

In future iterations, if additional experimental or simulated data regarding other properties of High Entropy Alloys (HEAs) becomes accessible, this methodology can be extended to predict those properties as well.In addition to predicting bulk modulus, future iterations of this project could explore the prediction of other mechanical properties such as ductility, toughness, and strength, as well as physical properties like thermal conductivity and electrical resistivity. By incorporating a wider range of properties, a more comprehensive understanding of HEAs could be achieved, enabling researchers to tailor their compositions for specific applications with greater precision. Furthermore, as the dataset grows and more advanced machine learning algorithms are developed, the accuracy and efficiency of the models could be further improved, making the development of HEAs even more accessible and impactful.

By incorporating a broader range of properties into the model, its accuracy and predictive power can be enhanced, leading to more reliable predictions for HEA development. This could potentially reduce the need for extensive computational simulations, making the process more cost-effective and efficient. Ultimately, this approach could revolutionize the way HEAs are designed and developed, paving the way for the creation of advanced materials with tailored properties for various engineering applications.

Chapter 8

Literatures cited/References

1. https://doi.org/10.1016/j.pmatsci.2022.101018
2. https://doi.org/10.1016/j.mtener.2021.100638
3. https://www.nature.com/articles/s41524-022-00779-7
4. https://www.mdpi.com/2075-4701/14/2/235#table_of_contents
5. **Code Availability** : https://github.com/yaswanth297/HEA_PropPrediction