Project No. 02313 / Grant No. JPMXP0723833151

Report form of Joint Research Project at ZAIKEN (FY2023)

Title of Project	Data analysis and machine learning based prediction of mechanical and physical properties for the development of energy saving materials		
Priority Area	I-B(Prolongation of service life & Structures), I-C(Prolongation of service life & Properties), III-B(Energy saving & Structures), III-C(Energy saving & Properties)		
Continuation of FY2022 project			
Name of Main Applicant	Nguyen Hai Chau		
Institution	University of Engineering and Technology, Vietnam National University, Hanoi	**title**	Associate Professor

Aim of the research project

Classical strategies to design new materials are often trial-and-error based and require many experiments and simulations, rendering them time-consuming and costly. In recent years, new methods, namely machine learning, are combined with orthodox methods in materials science, such as quantum chemistry and first principles calculations to speed up and reduce cost of material discovery process, utilizing data from already discovered materials. In this project, we aim to use statistical data analysis and machine learning methods to predict mechanical and physical properties such as hardness and yield strength of high-entropy alloys as energy saving materials. The potential contributions of this research include the collection, enrichment and analysis of high-entropy alloys databases with the aforementioned properties, as well as the development of machine learning prediction models for the hardness and yield strength of high-entropy alloys.

Contents and results of the research

In recent years, machine learning (ML) is increasingly integrated with traditional methods in materials science, such as quantum chemistry and first principles calculations, to predict novel materials with desired properties. However, most of the ML methods typically offer point predictions, i.e. single continuous (e.g. yield strength) or discrete (e.g. phase) values. Prediction regions such as intervals of sets of discrete values are often unavailable.

High-entropy alloys (HEAs) are alloys composed of five or more principal elements with near-equal concentrations. Recently, HEAs have garnered significant research attention due to their superior mechanical and physical properties, such as yield strength and hardness, in comparison to conventional alloys [1, 2]. ML is one of the main methods to predict properties of novel HEAs based on simulation and experimental data [3, 4]. ML methods like random forests [5] and deep learning [6] have demonstrated high accuracy in predicting such properties [7, 8]. However, the uncertainty of predictions, i.e. the prediction region, is often unavailable.

In FY2023 project, using conformal prediction methods [9], we quantified relationship between prediction regions size and size of a HEA's yield strength database along with strategies for training and validating ML models. The preliminary results revealed that on average, the size of prediction regions of leave-one-out is 49.83 MPa smaller than that of simple split training/validation strategy. Additionally, the size of prediction regions decreases by 3.18 Mpa if the size of training dataset increases by 1 [10].

References

1. D. B. Miracle, O. N. Senkov, Acta Mater. 122, 2017, 448-511.
2. Y.F. Ye, et al., Mater. Today 19(6), 2016, 349-362.
3. C. K. H. Borg, et. al., Sci Data 7, 2020, 430.
4. S. Gorsse, et al., Data in brief 21, 2018, 2664-2678.
5. L. Breiman, Random forests, Machine learning, vol. 45, 2091, 5-32.
6. B. Yoshua, Y. LeCun and G. Hinton, Nature, 521, 2015, 436-444.
7. U. Bhandari, et al., Crystals 2021, 11, 46.
8. U. Bhandari, et al., Yield strength prediction of high-entropy alloys using machine learning, Volume 26, 2021, 101871.
9. G. Shafer and V. Vovk. A tutorial on conformal prediction, Journal of Machine Learning Research, 9:371–421, 2008.
10. N. H. Chau, T. Yamamoto, Uncertainty quantification for prediction of high-entropy alloy's yield strength, 7th International Symposium on Frontiers in Materials Science (FMS 2024), National Yang Ming Chiao Tung University, Hsinchu, Taiwan, January 21st-24th, 2024.

Outputs of the project (publications, presentations, patents)

N. H. Chau, T. Yamamoto, Uncertainty quantification for prediction of high-entropy alloy's yield strength, 7th International Symposium on Frontiers in Materials Science (FMS 2024), National Yang Ming Chiao Tung University, Hsinchu, Taiwan, January 21st-24th, 2024.