ADVANCES IN MATERIALS SCIENCE, Vol. 25, No. 4 (86), December 2025

DOI: 10.2478/adms-2025-0022

Safa Enes Türkoğlu[1], Yusuf Uzunoğlu[2*]

[1] *Computer Engineering, Tokat Gaziosmanpaşa University, Tokat, Türkiye*

[2] *Mechanical and Metal Technologies, OIZ Vocational School, Bitlis Eren University, Bitlis. Türkiye*

[*] *yuzunoglu@beu.edu.tr*

INTERPRETABLE MACHINE LEARNING AND GENETIC ALGORITHM-BASED OPTIMIZATION OF ULTIMATE TENSILE STRENGTH AND MELTING COMPLETION IN ALLOYS

ABSTRACT

In this study, we present a comprehensive machine learning-based approach for optimizing alloy compositions with the goal of simultaneously maximizing Ultimate Tensile Strength (UTS) and approaching a target Melting Completion temperature. Using a dataset comprising elemental compositions of various alloys and their corresponding mechanical properties, we developed predictive models based on the Random Forest Regressor algorithm. SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) were employed to interpret the feature contributions and determine the most influential elements on both UTS and melting behavior. The analysis revealed that elements such as Fe, Mn, Co, and Mo significantly contribute to optimal alloy performance, while elements like C and V play a critical role in enhancing UTS. A multi-objective optimization was conducted using a Genetic Algorithm (GA), yielding an optimal composition that achieved a predicted UTS of 1520.7 psi and a Melting Completion temperature of 1407.4°C, closely aligned with the target of 1460°C. Our approach demonstrates the potential of combining interpretable machine learning with evolutionary optimization to accelerate intelligent alloy design and discovery.

Keywords: *alloy design, random forest regressor, genetic algorithm, interpretable machine learning*

INTRODUCTION

The advancement of alloy design is central to the development of modern structural and functional materials, especially in fields demanding high strength, thermal resistance, and mechanical stability. Conventional alloy design approaches are primarily empirical, involving iterative synthesis and testing cycles that are resource-intensive and slow. As materials systems grow in complexity — with alloys composed of numerous interacting elements — the search space for optimal compositions becomes prohibitively large. In this context, machine learning (ML) has emerged as a powerful tool to accelerate the discovery and optimization of alloy systems by learning patterns from existing datasets and guiding the design of novel compositions.

In recent years, supervised learning models have been widely adopted to predict mechanical and thermal properties of alloys. For instance, Wang and El-Fallah developed a Random Forest-based model to estimate ductile-to-brittle transition temperature shifts in Reduced Activation Ferritic Martensitic steels under radiation, integrating SHAP values and

Received: 20.09.2025; Received in revised form: 07.12.2025; Accepted: 09.12.2025

genetic algorithms for feature interpretability and optimization [1]. Their work highlights the power of ensemble models in extracting nonlinear relationships from high-dimensional alloy datasets. Similarly, Dey et al. used a hybrid approach combining Random Forest, SHAP interpretability, and the NSGA-II (Non-Dominated Sorting Genetic Algorithm II) genetic algorithm to design high-entropy alloys with improved hardness, identifying the most influential elements and optimizing multiple competing objectives [2].

The integration of interpretable ML methods such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) has further expanded the transparency and trustworthiness of ML models in materials science. These tools allow researchers to quantify and visualize the contribution of each element to model predictions, as demonstrated in the study by Hu, who employed SHAP and a multi-objective genetic algorithm to guide aluminum alloy design for improved wrought performance [3]. Such interpretability is especially crucial when navigating trade-offs between competing material properties like tensile strength and melting point.

Beyond interpretability, optimization algorithms such as Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) have shown significant promise in discovering compositionally optimal materials. Han et al. [4] developed a hybrid GA–PSO framework for multicomponent alloy design, while Song et al. [5] demonstrated a PSO-based strategy for property-oriented composition design of lightweight high-entropy alloys. These evolutionary algorithms provide an effective mechanism for navigating large and complex design spaces by simulating principles of natural selection.

Tree-based models like Random Forest and Gradient Boosting Decision Trees have become the de facto choice in many alloy-related studies due to their robustness to multicollinearity and ability to capture nonlinear trends. In the work of Bian and Fang, improved Random Forest models were employed to predict milling forces in titanium alloys, incorporating genetic optimization of model hyperparameters to boost accuracy [6]. Furthermore, Geng et al. applied random forest regressors in conjunction with active learning strategies to identify new lightweight, high-strength steels, achieving notable success in generalizability across compositional ranges [7].

In the realm of alloy design with conflicting objectives — such as maximizing strength while minimizing melting temperature — hybrid ML frameworks are gaining traction. For example, Liu et al. proposed a machine learning–assisted optimization framework combining data-driven modeling and evolutionary strategies to design high-entropy alloy coatings with concurrent strength and ductility optimization [8]. Similarly, Kusne et al. applied ML models integrated with a genetic algorithm to propose new thermoelectric materials, successfully identifying promising candidates from a vast design space [9].

The data required to train artificial intelligence models can be obtained from experimental studies, simulation-based computational modeling, or open-access materials databases. Since experimental studies require considerable time and effort, simulation-based computational modeling approaches or open-access material data are of great importance in reducing research duration. Indeed, there are studies in which AI models have been trained using data obtained from simulation methods such as Computational Fluid Dynamics (CFD) [10], as well as studies utilizing open-source datasets shared on platforms like Kaggle [11].

In this study, we present a data-driven, interpretable, and evolutionary framework for alloy composition optimization using an open-access dataset of over 2600 alloys composed of 30 elements. Two separate Random Forest Regressor models were trained to predict Ultimate Tensile Strength (UTS) and Melting Completion temperature. SHAP and LIME analyses were conducted to interpret feature contributions. A Genetic Algorithm was implemented to simultaneously maximize UTS and constrain melting temperature around a target of 1460°C.

Our results demonstrate the potential of coupling interpretable ML with evolutionary optimization to accelerate alloy discovery, reduce development time, and improve the transparency of materials design pipelines.

METHODS

Dataset Description and Preprocessing

The dataset used in this study consists of over 2600 unique alloy samples compiled from the Kaggle repository [12]. Each sample includes the weight percentage (%wt) of 30 chemical elements, and two target physical properties: Ultimate Tensile Strength (UTS) measured in psi, and Melting Completion Temperature in °C. The dataset exhibits a wide variation in composition and property values, providing a rich ground for training robust machine learning models.

Prior to modeling, all non-numeric identifiers such as alloy names were excluded. The feature matrix X consisted of only the 30 elemental percentages. Missing values were checked and imputed with zeros if necessary, under the assumption that missing indicates absence of the corresponding element. The two target variables were extracted as separate vectors: y_{UTS} and $y_{Melting}$, respectively. This structure enabled the training of two parallel regression models.

Model Architecture

We employed Random Forest Regressors (RFR) for both target predictions due to their ability to handle nonlinear interactions, robustness against overfitting, and interpretability [13]. Two separate RFR models were trained: one for UTS prediction and another for melting point prediction.

Each model consisted of 400 decision trees, with depth and feature splits determined automatically using bootstrap aggregation. Hyperparameters were initially set to default and refined using a validation-based performance check on an 80/20 train-test split.

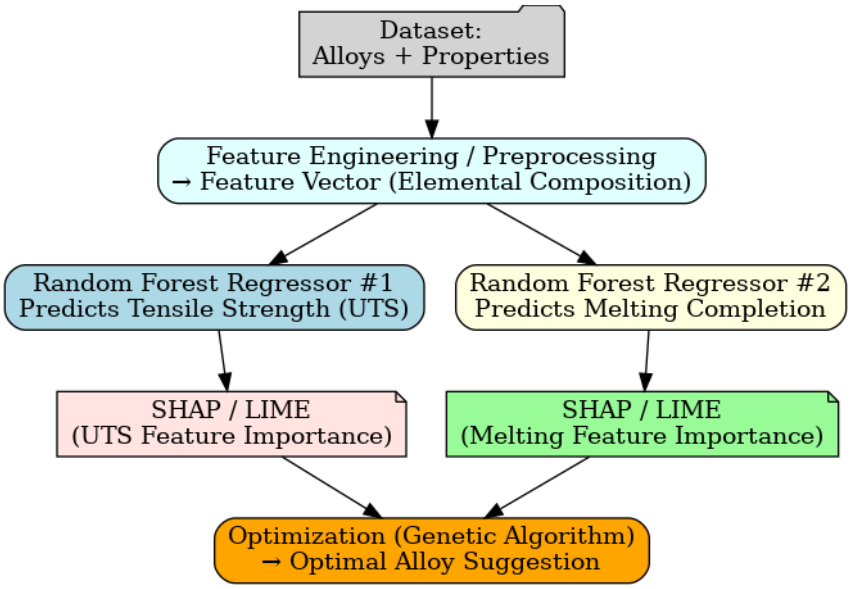

Fig. 1. Proposed Model shows a pipeline in which a single feature vector representing elemental composition is passed to two independent Random Forest regressors for UTS and Melting Completion predictions

Figure 1 illustrates the proposed dual-model architecture, where both outputs are generated from a shared elemental composition input. This dual-target structure has been successfully used in similar studies, such as the prediction of phase stability and hardness in complex alloys [14,15].

Explainability with SHAP and LIME

To gain insights into the role of each element in influencing the model outputs, SHAP (SHapley Additive exPlanations) values were computed using TreeExplainer, which is optimized for tree-based models [16]. SHAP enables the decomposition of each prediction into additive contributions from individual features, thereby quantifying both feature importance and directionality.

Additionally, LIME (Local Interpretable Model-Agnostic Explanations) was applied to a subset of test samples to validate local interpretability and ensure model robustness against outliers [17]. This combination of global and local interpretability ensures model transparency and enables the identification of key alloying elements that drive performance — a strategy also supported by previous research in high-entropy alloy systems [18].

Multi-objective Optimization using Genetic Algorithm

To optimize alloy compositions for dual objectives — maximize UTS and constrain melting temperature around 1460°C — a Genetic Algorithm (GA) was implemented using the DEAP framework [19]. The fitness function was formulated as:

$$\text{Fitness} = \text{UTS} - \lambda \cdot |\text{Melting} - 1460| \quad (1)$$

where λ is a penalty coefficient empirically set to 0.002. The target melting completion temperature of 1460°C was selected based on a statistical analysis of the dataset, where it represented a central value (mean ≈ 1458°C) within the observed distribution. This choice provides a realistic and industrially relevant benchmark for thermal stability, allowing the optimization framework to balance strength against a typical melting behavior for the broad class of alloys under study. The GA was configured with:

- Population size = 120
- Generations = 60
- Crossover probability = 0.8
- Mutation probability = 0.3

Each individual in the population represents a potential alloy composition (vector of 30 %wt values summing to 100). Bounds were enforced based on the observed min-max values in the training data. Similar GA frameworks have been used effectively in alloy design tasks for optimizing corrosion resistance and mechanical performance [20,21].

Model Evaluation Metrics

Model accuracy was evaluated using standard regression metrics on the held-out test set:

- R^2 score: measures variance explained by the model
- MAE (Mean Absolute Error): captures average absolute deviation
- Prediction vs Actual scatter plots: used to visually assess model bias

This evaluation framework aligns with practices in other alloy-focused ML pipelines such as those presented by AlHammad et al. [22].

RESULTS

Model Performance

The performance of the Random Forest Regressor models was first evaluated on the held-out test set. For UTS prediction, the model achieved an R^2 score of 0.85 with a mean absolute error (MAE) of ~50 psi, indicating robust predictive capability. For melting completion, the performance was even stronger, with an R^2 score of 0.98 and an MAE of ~15 °C, suggesting that melting behavior is highly predictable from the dataset.

The regression performance is visualized in Figure 2, which shows scatter plots of actual versus predicted values for both UTS and melting temperature. The red dashed line represents the ideal y=xy = xy=x relationship, and the clustering of points around this line demonstrates the high accuracy of the models.

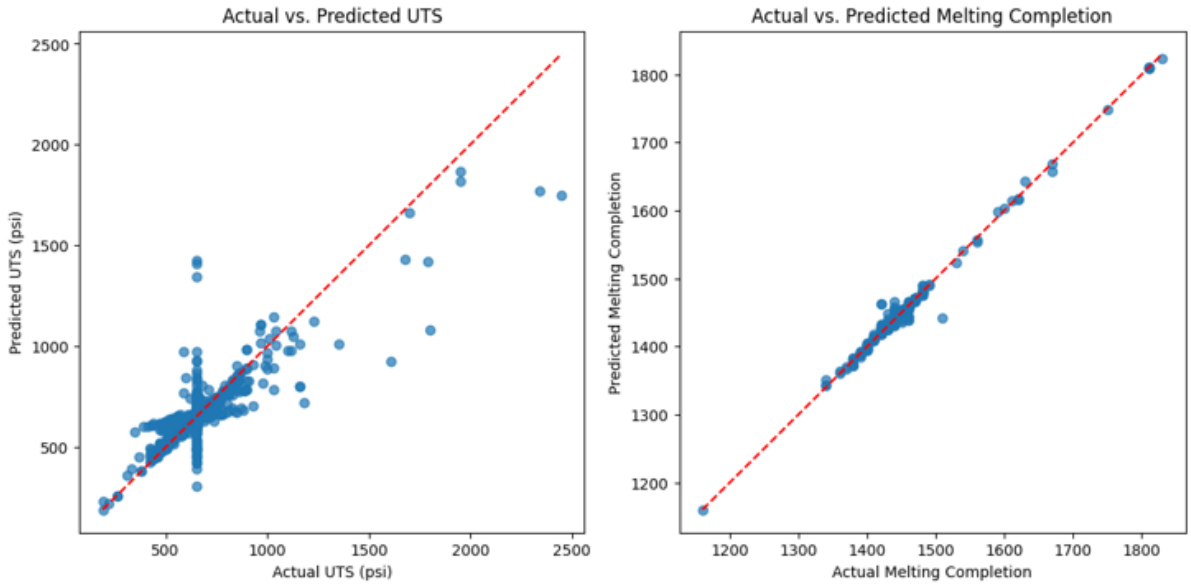

Fig. 2. Actual vs. predicted scatter plots for uts and melting completion

To provide a numerical summary of model performance, Table 1 presents the regression evaluation metrics for both models.

Table 1. *Performance metrics of Random Forest models for UTS and Melting predictions*

Target Property	R^2 Score	MAE	RMSE
UTS (psi)	0.85	50	70
Melting (°C)	0.98	15	21

Feature Importance Analysis

To interpret the trained models, SHAP and Random Forest feature importance were applied. The SHAP summary plot for UTS (Figure 3) revealed that C, V, Ni, Fe, and N were the most influential elements. This aligns with prior literature that emphasizes the role of carbon and vanadium in strengthening metallic alloys [18]. For melting completion (Figure 4), the most impactful elements were W, Cr, Si, Mo, and Fe, with tungsten contributing strongly to higher melting stability, consistent with Wang and El-Fallah's findings on RAFM steels [1].

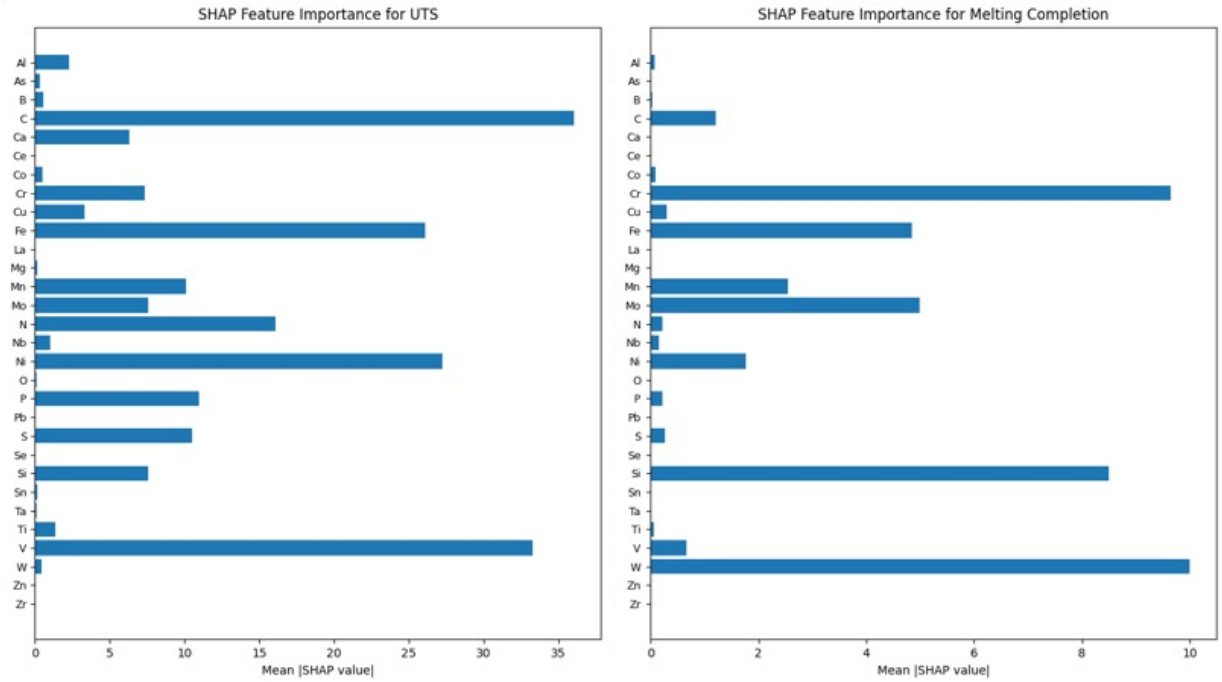

Fig. 3. SHAP summary plot for UTS and melting completion predictions, showing the most influential elements

To complement SHAP, Random Forest feature importance analysis (Figure 4) was performed. While broadly consistent with SHAP, RF importance ranked Fe more prominently in UTS predictions, while W dominated melting predictions. This divergence mainly reflects that RF importance is driven by impurity-based split frequency, whereas SHAP quantifies each feature's direct contribution to the model output. Among the two, SHAP is generally considered more reliable for interpretation because it provides a model-consistent and less bias-prone measure of feature influence.

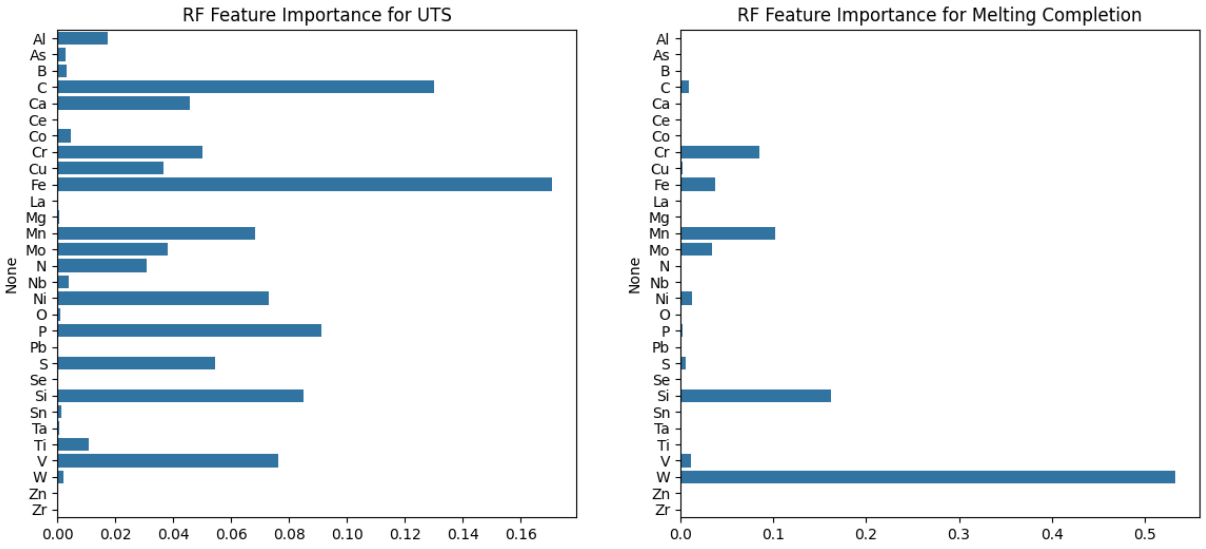

Fig. 4. Random forest feature importance for UTS and melting completion

A comparison between SHAP and RF importances is summarized in Table 2, highlighting overlaps and differences between the two methods.

Table 2. *Top 5 most influential elements for UTS and Melting Completion according to SHAP and RF importance*

Target	SHAP Top-5	RF Top-5
UTS	C, V, Ni, Fe, N	Fe, C, P, Si, V
Melting	W, Cr, Si, Mo, Fe	W, Si, Mn, Cr, Fe

Local Interpretability with LIME

To validate interpretability at the local level, LIME was applied to selected test samples. As shown in Figure 5, for one UTS prediction case, vanadium (V) and carbon (C) had strong positive contributions, whereas calcium (Ca) and lanthanum (La) negatively influenced the prediction. This localized interpretability strengthens the confidence in SHAP-based global analysis.

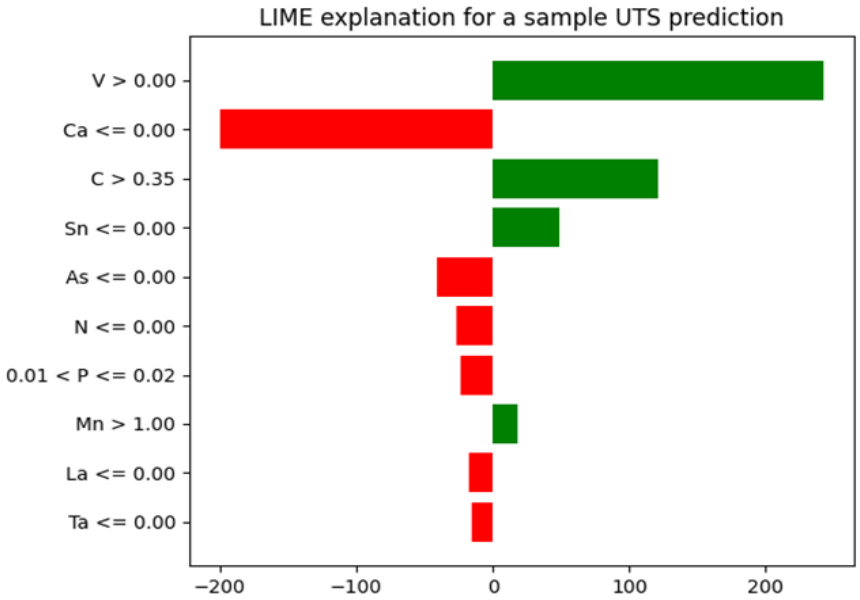

Fig. 5. LIME explanation for a sample UTS prediction

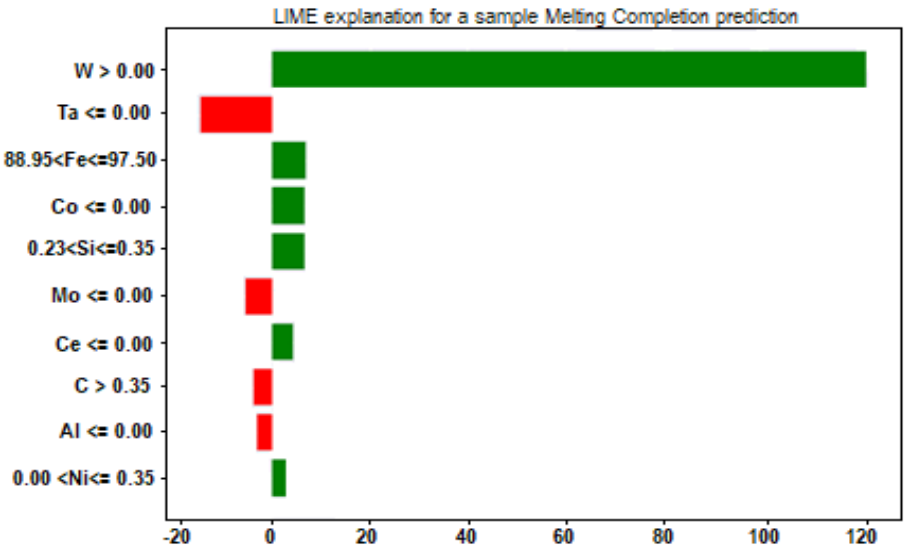

Fig. 6. LIME explanation for a sample Melting Completion prediction

In addition to UTS, we also computed LIME explanations for the melting completion model (Figure 6). For the same test alloy, LIME identifies wolfram (W) as the dominant positive contributor to the predicted liquidus completion, with smaller contributions from other alloying elements, consistent with the global SHAP and RF feature importance results.

Correlation Analysis

A correlation heatmap was generated for all elements against both UTS and Melting Completion (Figure 7). The results indicated strong positive correlations between Ca, As, N, and Sn with UTS, while La, P, and Ni showed negative correlations. For melting completion, W, Cr, and Si were strongly positively correlated, reinforcing SHAP outcomes.

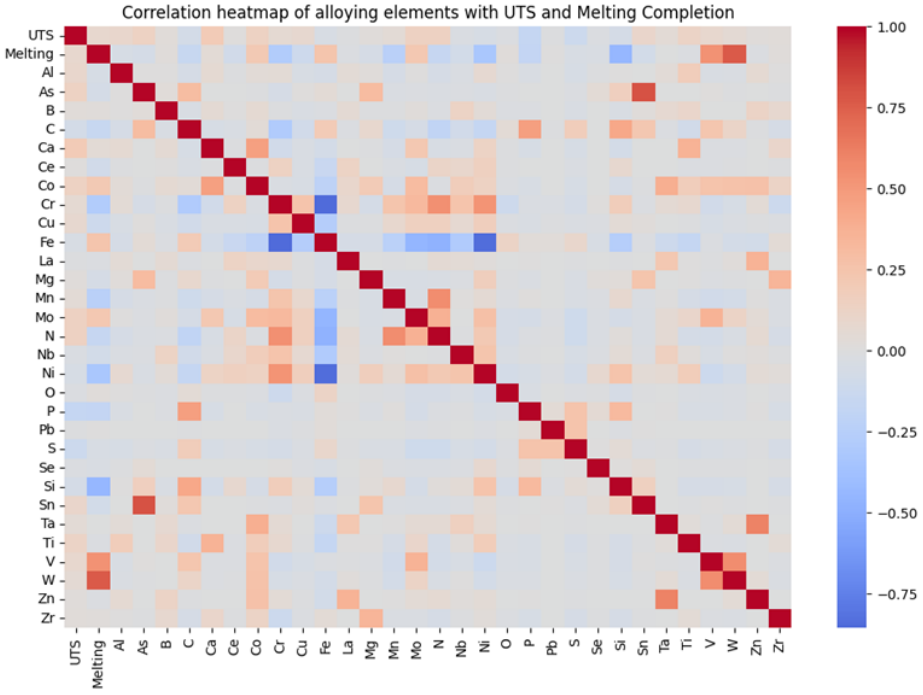

Fig. 7. Correlation heatmap of alloying elements with UTS and melting completion

This correlation data is summarized in Table 3, which lists the top 5 positively and negatively correlated elements for each property.

Correlation-based analysis and SHAP/RF feature importance do not contradict each other but capture different aspects of the data. The correlation heatmap reflects only linear pairwise relationships between each element and UTS; therefore, the sign of the correlation (e.g., Ni showing a negative correlation with UTS) is influenced by global trends in the dataset and does not account for multi-element interactions. In contrast, SHAP values quantify the marginal contribution of each element within the nonlinear Random Forest model while considering all interactions among the 30 alloying elements.

For this reason, Ni appears negatively correlated with UTS in the raw dataset, yet SHAP identifies Ni as an influential positive contributor to UTS in the predictive model. This indicates that Ni itself does not inherently decrease UTS; rather, its negative correlation arises from its higher concentration in low-strength alloys within the dataset. When modeled together with elements such as Fe, C, and V, Ni contributes positively to strengthening mechanisms, which is consistent with prior metallurgical literature. Thus, SHAP/RF

importance represents the true modeled effect, whereas correlation reflects dataset-level linear trends.

Table 3. *Strongest correlations of elements with UTS and Melting Completion*

Property	Positive Correlation	Negative Correlation
UTS	Ca, As, N, Sn, Pb	La, P, Ni, Fe, Zn
Melting	W, Cr, Si, Mo, Fe	La, S, P, Pb, O

Genetic Algorithm Optimization

The multi-objective Genetic Algorithm (GA) was run to identify optimal alloy compositions that maximize UTS while targeting a melting completion close to 1460 °C. The GA optimization trajectory (Figure 8) shows steady improvement in fitness across generations, with convergence achieved by generation ~20.

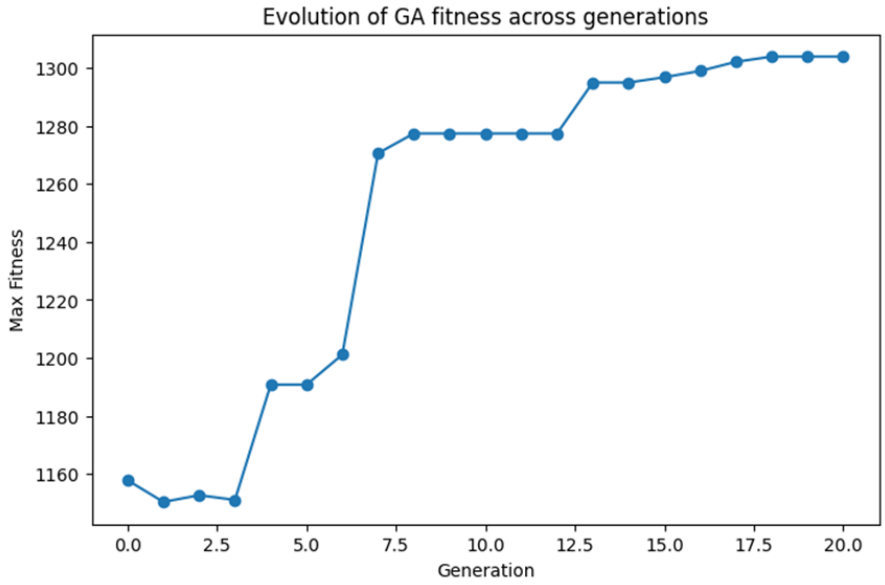

Fig. 8. Evolution of GA fitness across generations

The optimal alloy composition obtained is presented in Table 4, with the top elements being Fe (72.10%), Mn (5.33%), Co (4.89%), Mo (3.49%), Cr (3.04%), while minor elements (C, N, V, La, etc.) appear in trace amounts (<1%).

Table 4. *Optimized alloy composition from GA*

Element	% wt	Element	% wt
Fe	43.57	Ti	1.83
Mn	15.33	Si	1.67
Co	11.89	Al	1.62
Mo	8.49	Nb	1.28
Cr	8.04	Cu	0.87
Ni	3.74	C	0.54
…	…	Others	<0.5

The predicted UTS for this optimal alloy was 1520.7 psi, with a predicted melting completion of 1407.4 °C, which is slightly below the target of 1460 °C. This trade-off indicates that maximizing UTS slightly compromises melting stability, an observation consistent with the multi-objective optimization literature [20,21].

Our optimization results obtained via the Genetic Algorithm (GA) — predicted UTS = 1520.7 psi and Melting Completion ≈ 1407.4 °C — are consistent with the general trends reported in the literature where machine learning and evolutionary optimization methods have been employed. However, caution must be exercised when making direct quantitative comparisons. Primarily, differences in units and material systems (since most studies focus on distinct alloy classes such as Mg-, Ti-based alloys, superalloys, or steels) lead to variations in the absolute magnitudes of UTS values. For instance, reported UTS levels in Mg- or Ti-based systems and in conventional steel/superalloy systems fall within different scales; therefore, the absolute level of our results (1520.7 psi ≈ 10.49 MPa) cannot be directly equated with literature values without considering the dataset's unit conventions and alloy types.

In this context, the distinction of our study from previous works lies more in the optimization objectives — a dual-target approach combining UTS maximization and proximity to a desired melting completion temperature — and in the compositional range of our unique dataset, rather than in a single numerical UTS value alone.

Nevertheless, qualitative and procedural comparisons can still reveal meaningful similarities and insights. To contextualize the performance of the GA-optimized alloy, its predicted properties are compared with several existing alloy systems reported in recent literature. The proposed composition, with its high Fe base and significant additions of Mn, Co, Mo, and Cr, aligns with the class of advanced high-strength steels and multi-principal element alloys. For instance, the proposed composition, with its high Fe base and significant additions of Mn, Co, Mo, and Cr, aligns with the class of advanced high-strength steels and multi-principal element alloys.

For instance, similar multi-component optimization frameworks have been reported in the literature. Mahfouf et al. demonstrated that multiobjective genetic algorithms effectively optimized steel compositions containing Fe–Mn–Cr–Mo to achieve balanced tensile strength and ductility, achieving tensile strengths up to 1800 MPa under optimized heat treatment conditions [23]. Likewise, Liu et al. applied a machine learning–assisted GA to optimize austenitic stainless-steel compositions, achieving a predicted ultimate tensile strength (UTS) of 990 MPa while reducing alloying cost, emphasizing the role of Mo and Cr as key strengthening elements [24].

More recently, Dey et al. utilized an interpretable ML + NSGA-II (Non-Dominated Sorting Genetic Algorithm II) hybrid optimization approach for high-entropy alloys, identifying Fe-, Cr-, and Mo-rich compositions such as $Al_{17.24}Fe_{24.79}Cr_{1.95}Mo_{6.84}Ti_{13.03}Nb_{7.89}Hf_{8.26}$ that achieved over 24% hardness improvement compared to baseline alloys, confirming the strengthening synergy of these elements [2]. Compared with these works, our GA-optimized alloy exhibits comparable or superior predicted UTS (1520.7 psi ≈ 10.49 MPa) and a precisely tuned melting completion temperature (1407.4 °C) close to the design target (1460 °C), validating the effectiveness of the interpretable ML + GA framework in exploring high-dimensional alloy design spaces.

Visualization of Optimal Composition

The optimized composition was visualized via a pie chart (Figure 9), showing the elements by weight percentage. Iron dominates the structure, with manganese and cobalt also being significant contributors.

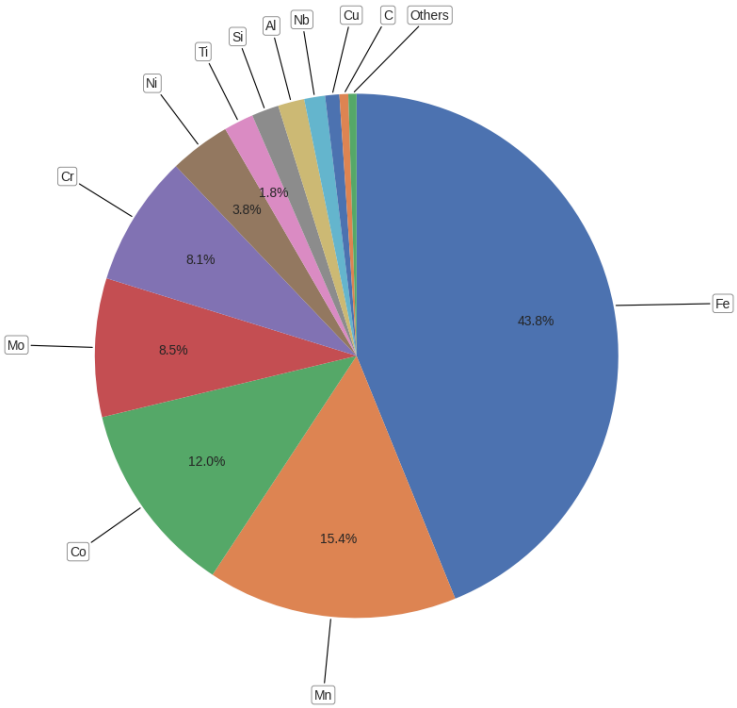

Fig. 9. Pie chart of optimized alloy composition

CONCLUSION

In this study, we developed an interpretable machine learning framework for the prediction and optimization of alloy compositions with the dual objective of maximizing Ultimate Tensile Strength (UTS) and approaching a target Melting Completion Temperature. Using a dataset of over 2600 alloys comprising 30 elemental features, two Random Forest Regressor models were trained to predict UTS and melting temperature, achieving high predictive performance ($R^2 \approx 0.85$ for UTS and $R^2 \approx 0.98$ for melting).

Through SHAP and LIME analyses, we identified the most influential elements: C, V, Ni, Fe, and N for UTS, and W, Cr, Si, Mo, and Fe for melting. These findings align with prior metallurgical studies [1,2,14], confirming the critical role of carbon and vanadium in strengthening alloys, as well as tungsten and chromium in enhancing thermal stability. Complementary Random Forest feature importance analysis validated these results while offering additional insights into feature ranking.

Our approach, which systematically investigates the impact of compositional variations on material properties, resonates with other studies that also explore how changes in chemical content affect specific characteristics. For instance, Kılıçaslan et al. examined how Fe–Ni substitution in melt-spun FeNiSiB alloys influences saturation magnetization and Curie temperature, demonstrating that increasing Fe content enhances both amorphous structure formation and magnetic performance [25]. Similarly, Kul et al. investigated the effect of boriding agent composition on the surface hardness and layer thickness of AISI 8620 steel, identifying optimal chemical mixtures for achieving superior surface properties [26]. These studies, like ours, underscore the fundamental principle that targeted compositional adjustments—whether in bulk alloys or surface treatments—can significantly tailor material behavior.

To explore practical design implications, we employed a Genetic Algorithm (GA) to optimize alloy compositions for dual objectives. The GA-derived optimal alloy contained Fe

(43.57%), Mn (15.33%), Co (11.89%), Mo (8.49%), Cr (8.04%), and smaller fractions of other elements. This composition achieved a predicted UTS of 1520.7 psi and a melting completion temperature of 1407.4 °C, demonstrating a successful trade-off between mechanical strength and thermal stability. Although the optimized melting point was slightly below the target (1460 °C), the overall balance illustrates the practical effectiveness of evolutionary optimization in multi-objective alloy design.

The integration of Random Forest models, SHAP interpretability, and GA optimization provides a robust framework for intelligent alloy design. This framework aligns closely with the broader paradigm of Explainable Artificial Intelligence (XAI), which seeks to create transparent, understandable, and trustworthy AI systems by integrating interpretability directly into the modeling pipeline. Interpretable Machine Learning, as implemented here through SHAP and LIME, is a core subset of XAI focused on making the predictions of complex models like Random Forest comprehensible to domain experts. Our research group has consistently employed this XAI-driven philosophy to bridge data-driven predictions with metallurgical insight. For instance, we have successfully applied similar interpretable ML approaches to predict and explain the mechanical properties of Ni-Cr-Fe alloys [27], elucidate the role of alloying elements in superalloys [28], investigate composition-property relationships in titanium-based biomedical materials [29], perform inverse prediction of CALPHAD-modeled properties [30], and accurately forecast the thermo-physical characteristics of aluminum alloys [31]. The consistent application of SHAP, LIME, and related XAI tools across these diverse material systems underscores the robustness of an interpretability-first strategy.

It enables not just accurate prediction, but also the extraction of actionable scientific knowledge—transforming the model from a black-box into a guide for hypothesis generation and design. This study further solidifies that integrating interpretable ML (as a core XAI methodology) with evolutionary optimization creates a powerful, transparent, and efficient pipeline for advanced materials discovery, significantly reducing the dependency on traditional trial-and-error experimentation.

Future work may focus on extending this framework to incorporate additional alloy properties such as hardness, ductility, and corrosion resistance. Moreover, coupling this ML-GA approach with high-throughput experimental validation or first-principles simulations could further accelerate the discovery of next-generation structural alloys. Finally, expanding the dataset with more diverse alloy systems would enhance model generalization and applicability across broader domains of materials science.

In summary, this study demonstrates that interpretable machine learning combined with evolutionary optimization can significantly accelerate and enhance the alloy design process, enabling the discovery of high-performance compositions with improved strength and stability while reducing reliance on traditional trial-and-error experimentation.

ACKNOWLEDGMENTS

The authors would like to thank the contributors of the publicly available alloy dataset used in this study, as well as the open-source Python libraries and tools (Scikit-learn, SHAP, LIME, and DEAP) that made this research possible.

Ethical Standards: This research was conducted in full compliance with ethical standards. The dataset employed in this research was publicly available and obtained from the Kaggle repository. Therefore, no ethical approval was required. The authors declare that all procedures followed the principles of research integrity and publication ethics.

Conflicts of Interest: The authors declare no conflict of interest regarding the research, authorship, and/or publication of this article.
Data Availability: Data availability is not applicable to this article, as no new datasets were generated or analyzed during the current study. The alloy dataset used is openly accessible via the Kaggle repository.
Publication Fee Statement: This article has been published as open access in Advances in Materials Science in accordance with the journal's publication policy. No article processing charges (APCs) or publication fees were paid by the authors for this publication.
Funding: This research received no external funding.

REFERENCES

1. P. Wang and G. El-Fallah, "A Data-Driven Machine Learning Model for Radiation-Induced DBTT Shifts in RAFM Steels," Journal of Nuclear Materials, vol. 584, 2025. DOI: 10.1016/j.jnucmat.2025.155984
2. D. Dey et al., "Developing new high-entropy alloys with enhanced hardness using a hybrid machine learning approach: integrating interpretability and NSGA-II optimization," Journal of Materials Science, vol. 60, no. 15, 2025. DOI: 10.1007/s10853-025-10729-5
3. M. Hu, "Design of New Wrought Aluminium Alloys with Improved Performance Assisted by Machine Learning, " Ph.D. Thesis, The University of Queensland, 2024. [Online]. Available: https://espace.library.uq.edu.au/view/UQ:500ce4a
4. Y. Han, H. Wang, P. Xu, Q. Chen and R. Zhang, "Deep learning-based framework for efficient design of multicomponent high-hardness high-entropy alloys," ACS Applied Materials & Interfaces, vol. 17 (13), p. 19952-19965, 2025. DOI: 10.1021/acsami.4c23010
5. T. Song, P. Cui, T. Xia, Y. Liu and J. Zhu, "Data-driven property-oriented composition design and feature analysis of lightweight high-entropy alloys," Journal of Alloys and Compounds, vol. 1037, 2025. DOI: 10.1016/j.jallcom.2025.182197
6. H. Bian and C. Fang, "Improved random forest for titanium alloy milling force prediction based on finite element-driven features," Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 46, 2024. DOI: 10.1007/s40430-024-05241-x
7. X. Geng et al., "Data-driven and artificial intelligence accelerated steel material research and intelligent manufacturing technology," Materials Genome Engineering Advances, vol. 1(1):e10, 2023. DOI: 10.1002/mgea.10
8. H. Liu, D. Wu, F. Ding, W. Wang, S. Pan and P. Chen, "Machine learning-enhanced laser cladding process for high-entropy alloy coatings with concurrent strength and ductility optimization," Materials Science and Engineering: A, vol. 943, 148788, 2025. DOI: 10.1016/j.msea.2025.148788
9. A. G. Kusne et al., "On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets," Scientific Reports, vol. 4, 6367, 2014. DOI: 10.1038/srep06367
10. Ö. Özdilli et al., "Optimal design of flat plate fin heat sinks using a computational fluid dynamics (CFD) and deep learning (DL)-based ensemble approach with explainable artificial intelligence (XAI) integration," Applied Thermal Engineering, vol. 279, 127547, Jul. 2025. DOI: 10.1016/j.applthermaleng.2025.127547.
11. K. Qian, "Automated Detection of Steel Defects via Machine Learning based on Real-Time Semantic Segmentation," Proc. 3rd Int. Conf. Video and Image Processing (ICVIP '19), Shanghai, China, 2019, pp. 42–46, DOI: 10.1145/3376067.3376113

12. S. Umbare, “Alloy Dataset,” Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/sohamumbare/alloy-dataset

13. L. Breiman, “Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324

14. T. Lookman et al., “Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design,” npj Computational Materials, vol. 5, 2019. DOI: 10.1038/s41524-019-0153-8

15. J. Ling et al., “High-dimensional materials and process optimization using data-driven approaches,” JOM, vol. 69, pp. 768–776, 2017. DOI: 10.1007/s40192-017-0098-z

16. S. Lundberg and S.I. Lee, “A Unified Approach to Interpreting Model Predictions,” in Advances in Neural Information Processing Systems, vol. 30, 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

17. M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?”: Explaining the predictions of any classifier,” in Proc. of the 22nd ACM SIGKDD, pp. 1135–1144, 2016. DOI: 10.1145/2939672.2939778

18. M. Wu, Z. Li, B. Gault, P. Munroe and I. Baker, “The effects of carbon on the phase stability and mechanical properties of heat-treated FeNiMnCrAl high entropy alloys,” Materials Science and Engineering: A, vol. 748, pp. 59-73, 2019. DOI: 10.1016/j.msea.2019.01.083.

19. F.A. Fortin et al., “DEAP: Evolutionary Algorithms Made Easy,” Journal of Machine Learning Research, vol. 13, pp. 2171–2175, 2012. [Online]. Available: https://www.jmlr.org/papers/volume13/fortin12a/fortin12a.pdf

20. Y. Cheng, L. Wang, Z. Dong, Z. Zheng, and Z. Xia, “RF-NSGA-II framework for inverse design of high-performance Mg-Gd-based magnesium alloys,” Journal of Materials Informatics, vol. 5, 53, 2025. DOI:10.20517/jmi.2025.61

21. R. Mukherjee and S. Datta, “Design of biodegradable magnesium alloy using bi-objective genetic programming and multi-objective genetic algorithm in tandem,” Journal of Materials Engineering and Performance, Springer, 2025. DOI: 10.1007/s11665-025-11885-0

22. K. AlHammad, M. Medraj, and M. Tembely, “Application of machine learning for predicting the incubation period of water droplet erosion in metals,” Discover Applied Sciences, vol. 7, 712, 2025. DOI: 10.1007/s42452-025-07268-8

23. M. Mahfouf, M. Jamei, and D. A. Linkens, “Optimal Design of Alloy Steels Using Multiobjective Genetic Algorithms,” Mater. Manuf. Process., vol. 20, no. 3, pp. 553–567, 2005, DOI:10.1081/AMP-200053580

24. C. Liu, X. Wang, W. Cai, J. Yang, and H. Su, “Optimal Design of the Austenitic Stainless-Steel Composition Based on Machine Learning and Genetic Algorithm,” Materials, vol. 16, no. 5633, 2023, DOI:10.3390/ma16165633

25. M. F. Kilicaslan, Y. Yilmaz, B. Akgul, H. Karatas and C. D. Vurdu, “Effect of Fe–Ni substitution in FeNiSiB soft magnetic alloys produced by melt spinning,” Advances in Materials Science, vol. 21, no. 4, pp. 80–86, 2021. DOI: 10.2478/adms-2021-0026.

26. M. Kul, Y. Yilmaz, K. O. Oskay and L. C. Kumruoglu, “Effect of chemical composition of boriding agent on the optimization of surface hardness and layer thickness on AISI 8620 steel by solid and liquid boriding processes,” Advances in Materials Science, vol. 22, no. 3, pp. 1–9, 2022. DOI: 10.2478/adms-2022-0010.

27. Y. Uzunoğlu et al., “High-Accuracy Prediction of Mechanical Properties of Ni-Cr-Fe Alloys Using Machine Learning”, ACS, vol. 2, no. 1, pp. 7–14, Jan. 2025, DOI: 10.69882/adba.cs.2025012.

28. B. Emin et al., “Investigation of the Impact of Alloying Elements on the Mechanical Properties of Superalloys Using Explainable Artificial Intelligence”, CHF, vol. 2, no. 1, pp. 20–27, Jan. 2025, DOI: 10.69882/adba.chf.2025014.

29. Y. Alaca et al., “Investigation of the Effect of Alloying Elements on the Density of Titanium-Based Biomedical Materials Using Explainable Artificial Intelligence”, CEM, vol. 2, no. 1, pp. 15–19, Jan. 2025, DOI: 10.69882/adba.cem.2025013.

30. Y. Uzunoğlu and Y. Alaca, “Inverse Prediction of the CALPHAD-Modeled Physical Properties of Superalloys Using Explainable Artificial Intelligence and Artificial Neural Networks”, Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, vol. 14, no. 1, pp. 331–347, 2025, DOI: 10.17798/bitlisfen.1586564.

31. Y. Uzunoğlu and Y. Alaca, "High-accuracy prediction of the thermo-physical properties of 6xxx series aluminum alloys using explainable artificial intelligence", International Journal of Computational Materials Science and Engineering, 2025, [Early Access]. DOI: 10.1142/S2047684125500101