Croatian Society for Materials Protection

Metalurgija/Metallurgy

Vol. 66 No. 1 / 2027

Prediction of Stability of Austenitic Doping Systems Based on First-Principles Calculation and Machine Learning

X.F. Jiao [1], S.H. Wang [1,3], Y.B. Song [1], R. Xie [1], K. Liu [2,3], and G.L. Ni [2,3,*]

https://doi.org/10.64486/m.66.1.1

[1] School of Metallurgy and Energy, North China University of Science and Technology, Tangshan, China
[2] Hebei Iron and Steel Laboratory, North China University of Science and Technology, Tangshan, China
[3] Tangshan Key Laboratory of Special Metallurgy and Material Manufacture, Tangshan, China
* Correspondence: e-mail: ngl@ncst.edu.cn

Type of the Paper: Article
Received: December 6, 2025
Accepted: May 9, 2026

Abstract: This study presents a novel methodology that combines machine learning with first-principles calculations to efficiently screen austenitic stabilizing elements. Correlation analysis identified the lattice constant as a critical factor influencing the stability of austenitic doping systems. Subsequently, the random forest, support vector regression, and AdaBoost models were evaluated, among which the random forest achieved the highest prediction accuracy (R^2 = 0.748). Furthermore, SHapley Additive exPlanations was employed to interpret the model, further verifying the potential role of the lattice constant. Based on the above studies, Cr, Ni, and Mn were identified as doping elements. Finally, the first-principles calculation is employed to verify the prediction results of the machine learning. First-principles calculation results revealed that the austenite doped with Cr exhibits the lowest system energy (-29609.59 eV) and solid solution energy (-11.795 eV). Electronic structure analysis (including charge density difference and density of states) reveals the underlying mechanism: the larger the lattice constant of the doping atom, the weaker its interaction with the iron atoms, as specifically manifested by the reduction in electron cloud density and covalent bonding. These findings not only confirm the scientific validity of the proposed integrated machine learning and first-principles approach but also offer important guidance for the rational design of high-performance austenitic materials.

Keywords: austenitic stainless steels; machine learning; doping; first-principles

1. Introduction

As a type of engineering metal material, austenitic stainless steels have received widespread attention in various fields, such as aerospace, biomedical, and nuclear reactors, due to their excellent wear and corrosion resistance as well as superior hot workability [1]. However, higher performance demands have been put forward for practical applications in special fields, which motivates researchers to explore new approaches to tune and obtain the desired microstructures and properties.

It is widely known that alloying elements are extremely important in affecting the performance of austenitic stainless steels, benefiting from the positive influence on both phase stability and mechanical behavior. Doping alloying elements not only can enhance the stability of austenitic phase, but also modulates the interactions between austenitic and second phases, which can further improve final mechanical properties [2,3].

Morales et al. [4] demonstrated that alumina nanoparticles doping can progressively enhance the corrosion resistance of austenitic stainless steel. Furthermore, the investigation on silver-doping in austenitic stainless steel by Yang et al. [5] demonstrates a nearly complete antibacterial efficacy (100 % against Escherichia coli, 99.5 % against Staphylococcus aureus), which highlights its biomedical potentiality. Whether alloying elements are doped in the form of solid solution or second phase, the microstructure can be refined by inhibiting recrystallization and grain growth to achieve fine-grained strengthening and precipitation strengthening.

Despite the proven efficacy of doping, the traditional trial-and-error experimental approach for identifying optimal alloying elements is often time-consuming and resource-intensive. This challenge underscores the need for more efficient and predictive material design strategies. Fortunately, recent advances of materials and computational science accelerate the combination of first-principles calculation and machine learning (ML) for materials development. With the combination of the two methods, it becomes possible to theoretically design materials with specialized properties, meanwhile, significantly reducing the development costs. Jiang et al. [6] predict the high-temperature fatigue life of 316 steel under physics-informed constraints via machine learning, which achieves the higher precision with minimum experiment. Furthermore, Zhang et al. [7] utilize density functional theory (DFT) based first-principles calculation to assess molybdenum doping effects on steel properties, which establishes a theoretical foundation for subsequent experimental validation. The above research illustrates the importance and practicability for developing materials via integrating machine learning with first-principles calculation. Nevertheless, current research has several limitations. Traditional trial-and-error experiments are costly and inefficient, whereas most theoretical or machine learning studies tend to be conducted independently, lacking effective integration between data-driven analysis and physical mechanism validation. Moreover, the key intrinsic factors governing the stability of austenitic doping systems remain ambiguous, and quantitative identification and in-depth mechanistic explanation are still insufficient.

Accordingly, this study proposes a novel methodology combining machine learning models with first-principles calculations to efficiently screen and understand the impact of doping elements on austenitic stability. First, a framework integrating correlation analysis, machine learning models, and SHAP analysis is established to screen for elements capable of stabilizing the austenite phase. Subsequently, first-principles calculations are employed to validate the model predictions by calculating the energy, solid solution energy, and electronic properties of the system, thereby establishing a new, efficient, and reliable methodology for material development and rational design.

2. Materials and Methods

2.1. Dataset Analysis

The feature vectors are derived from data published in academic journals [8], while the target vector is calculated using the CASTEP module. As shown in Figure 1, the feature vectors are constructed using the intrinsic properties of the doping elements, including atomic number, atomic radius, lattice constant, and standard enthalpy of formation. This target vector, which represents the system energy of doping austenitic systems and serves as a direct indicator of stability, is given in units of eV.

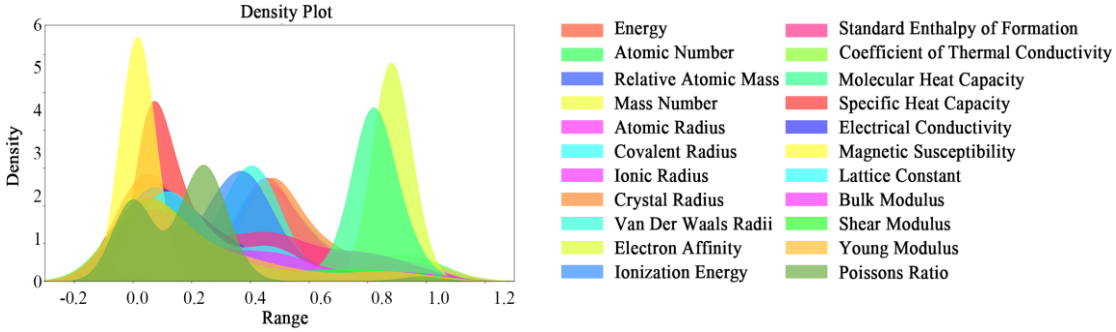


Figure 1. Ridge plot of feature values.

https://doi.org/10.64486/m.66.1.1

Based on the aforementioned features, a dataset comprising 110 doping austenitic systems is constructed. Although the sample size is limited, studies indicate that small datasets offer unique advantages in computational efficiency, model interpretability, and applicability to specific scenarios [9,10]. Subsequently, the dataset is normalized to mitigate scale disparities among features, ensure equal contribution of all variables to the model, and approximate a Gaussian distribution, which is a property preferred by many machine learning models [11]. As shown in Figure 1, the vast majority of features exhibit a Gaussian distribution after normalization [12]. Notably, although normalization makes most features follow a Gaussian distribution, a subset retains non-Gaussian residuals. Therefore, to mitigate potential biases, the model selection and evaluation metrics are rigorously optimized to ensure effectiveness and robustness [13].

2.2. Correlation Analysis

To address the challenges of multivariate optimization, this study employs Pearson correlation coefficient analysis to systematically evaluate the correlation between the feature vector and the target vector [14]. System energy is selected as the key indicator. This process enables the identification and removal of weakly correlated features, thereby enhancing model interpretability and predictive accuracy.

Figure 2 presents a heatmap of the Pearson correlation coefficients (R) among the diverse features, highlighting the linear relationships critical to the stability of the doping system. It can be clearly observed that the lattice constant of doping elements emerges as the most influential factor, with an R value of 0.308, exhibiting the highest linear correlation with system energy. Statistically, this corresponds to a weak linear relationship with $R^2 \approx 0.09$, which means that only about 9% of the variance in system energy can be explained by the linear correlation with lattice constant. It should be emphasized that this correlation analysis serves only as an initial feature screening step, not as a standalone predictive basis. Consequently, the lattice constants of doping elements are a key priority when selecting candidate dopants. Relevant studies indicate that the atomic radius of doping elements directly influences the degree of lattice distortion in austenite steel [15,16], excessive lattice distortion leads to a reduction in crystal structural stability. It is worth noting that the atomic radius contributes significantly to the variation in lattice constants. Therefore, the lattice constant becomes a key factor affecting the stability of the doping system. As discussed above, crystal stability is influenced by doping concentration, atomic radius and lattice constant, which collectively induce lattice distortion and thereby compromise stability [17-19]. In addition, shear modulus (R = -0.31), Young's modulus (R = -0.360), and enthalpy of formation (R = -0.265) show notable negative correlations with the energy of the austenite system. Thus, SHAP analysis was adopted to further clarify the influence of doping element parameters on austenite stability. Meanwhile, to enhance model generalizability, features with negligible correlations ($R \approx 0$) are systematically excluded during model training [20].

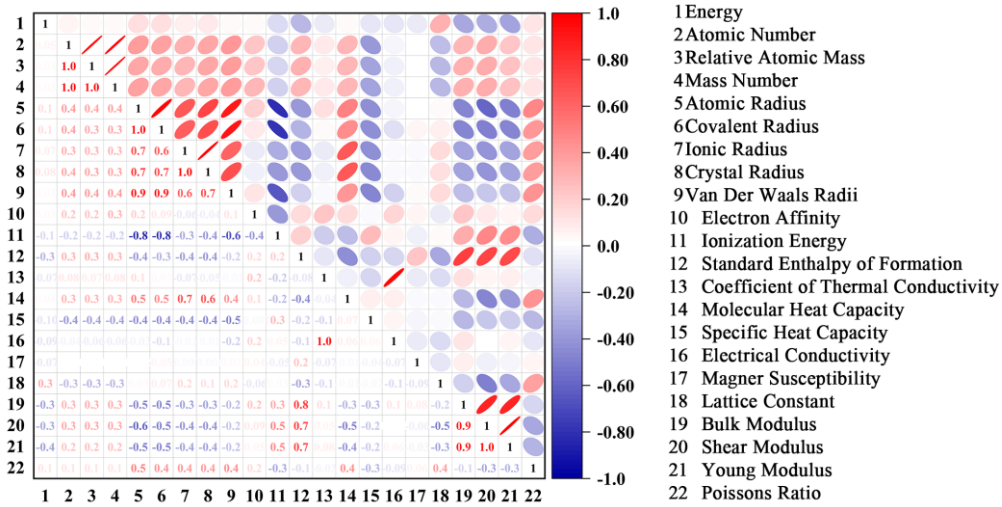


Figure 2. Heatmap of Pearson correlation coefficient (R values).

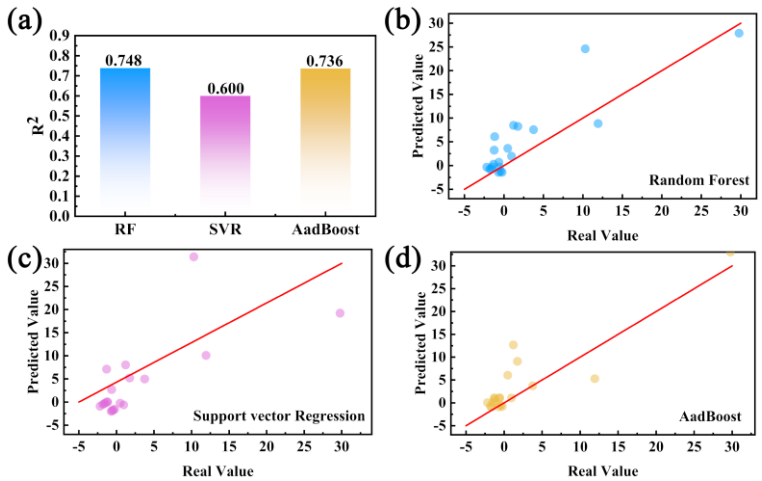
https://doi.org/10.64486/m.66.1.1

3. Results and Discussion

3.1. Model Analysis

To quantitatively evaluate the comprehensive impact of lattice constants and other input features in the process of predicting austenite stability using machine learning models, this study evaluates each algorithm model comprehensively by calculating the coefficient of determination (R^2) on an independent test set. Subsequently, the SHAP values of the optimal model on the test set are adopted to conduct an interpretable analysis of the algorithm model and quantify the contribution of each input feature to the model prediction results. Finally, by combining correlation analysis, model evaluation, and SHAP interpretation, the influence of input features on system stability is comprehensively assessed, and three candidate elements for austenite doping are screened out.

Figure 3(a) shows the average 10-fold cross-validation scores of each algorithm. It can be seen that the random forest (RF) model performs best. The R^2 value of the RF model is 0.748, which already demonstrates practicality and effectiveness [21]. Its number of decision trees is 100, the maximum depth is 10, and the maximum number of leaf nodes is set to 30. To further validate the model, the prediction behavior was examined in more detail: no significant difference in prediction error distribution is observed between the training and test sets, and the results of multiple cross-validation runs are stable, suggesting that the model does not suffer from obvious overfitting and exhibits good generalizability despite the small dataset. Next, the algorithm with excellent performance is AdaBoost, where the number of weak learners used in its regressor is 10. For the support vector regression (SVR) algorithm, both the penalty factor C and epsilon are set to 0.01, the kernel function is the radial basis function, and the gamma parameter is set to 0.1. Figure 3(b-d) compare predicted values with the actual values across various models. Overall, the correspondence between the predicted values and the actual values for the three algorithms is mainly concentrated in the low energy range of -2.5 to 0 eV. This finding is consistent with the previous analysis of the dataset and further confirms the accuracy of the model predictions. Moreover, the RF model demonstrates the strongest alignment, as evidenced by the $y = x$ regression line. Hence, it is selected as the best performing model among all machine learning prediction models evaluated in this study. In addition, the R^2 value of the SVR algorithm is lower than the other two algorithms, indicating that among the three models, SVR performs relatively poorly on the prediction task. The reason is that SVR is a margin-based regression method that seeks to find a regression hyperplane as smooth as possible while keeping the prediction error within a certain margin [22]. The focus of the machine learning part is to identify the key features that have the greatest impact on system stability. Through the dual validation of correlation analysis and subsequent SHAP analysis, the key features affecting stability have been clearly identified.

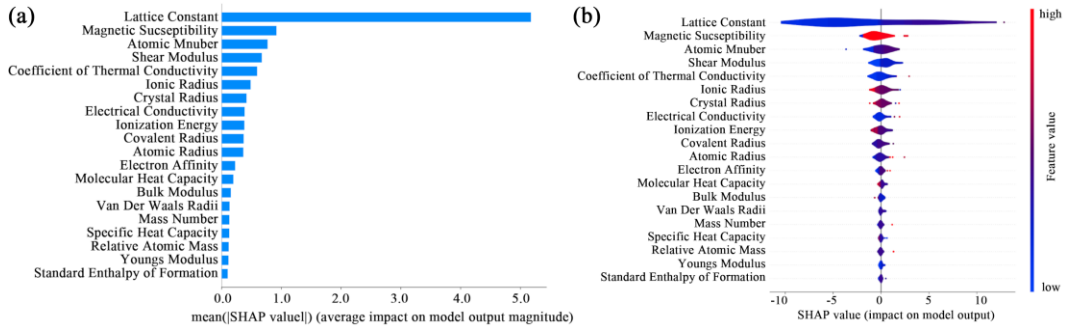


(a) Mean R^2 scores of three algorithms; (b-d) Correlation plots between predicted and actual values for (b) RF, (c) SVR, and (d) AdaBoost.

Figure 3. Performance evaluation of Machine learning algorithm models.

https://doi.org/10.64486/m.66.1.1

Figure 4 shows the SHAP analysis results for each feature in the RF model after correlation screening. Figure 4(a) is a distribution map of feature importance, which is a tool to visually illustrate the degree of influence of each feature on the model prediction. The initial Pearson correlation analysis and literature shows that the lattice constants are strongly correlated with the stability of doping systems. To further validate this finding, SHAP values are used to quantify the contribution of each feature to the model output [23]. It is found that the lattice constants exhibit a significantly higher mean SHAP value than other features, which proves it pivotal role in the stability of doping systems. Furthermore, the magnetic susceptibility of elements also has a significant influence. The higher the magnetic susceptibility of the doping element, the more stable the austenite structure after doping [24]. Figure 4(b) shows the violin plot of the distribution of each feature value. It can be seen that the color distribution of the lattice constant feature is relatively darker, indicating that it has a deeper influence on the predictions of the RF model. Meanwhile, the violin plot exhibits a wide lateral shape, indicating that the corresponding feature is broadly distributed across the dataset and contains diverse values [25].



(a) overall feature map of SHAP features; (b) violin plot of importance values of SHAP features.

Figure 4. SHAP analysis of RF model.

Several limitations and uncertainties should be noted. First, the input features are restricted to intrinsic atomic properties (e.g., atomic radius, lattice constant, enthalpy of formation), while external factors such as doping concentration, temperature, and pressure are not considered; therefore, the model is applicable only to similar doping systems. Second, as a black-box model, the random forest cannot directly output prediction intervals; the uncertainty mainly arises from the small sample size, potential noise in literature-derived feature values, and the inherent stochasticity of the algorithm.

Based on the above analysis, using the lattice constant as a key screening criterion, three elements with different lattice constant values were selected, as presented in Table 1. The relationship among the length units is as follows: 1 nm = 10 Å = 1000 pm. It can be found that the selected elements Cr, Ni, and Mn have atomic radius similar to that of Fe, which can effectively control the degree of lattice distortion in austenite [15,16]. Subsequently, calculations were performed on austenite doped with different elements to validate the predictions of the ML model.

Table 1. The atomic radius and lattice constant of the doping elements.

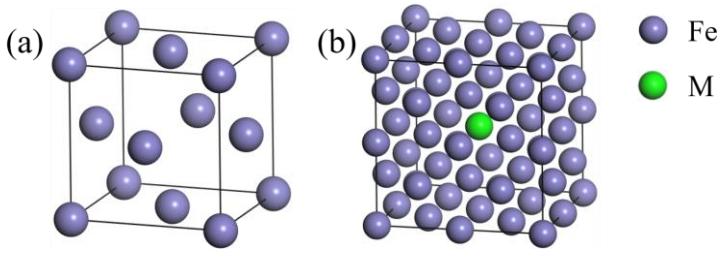
Doping elements	Atomic radius/pm	Lattice constant/Å
Cr	140	2.88
Ni	135	3.52
Mn	140	8.89

3.2. Crystal Structure Models

First-principles calculations were performed using the CASTEP module based on density functional theory (DFT) to investigate the stability and electronic properties of austenite doped with Cr, Ni, and Mn [26,27]. To verify the machine learning predictions, the PBE functional under the generalized gradient approximation

https://doi.org/10.64486/m.66.1.1

(GGA) was used to treat the exchange-correlation energy, thereby obtaining the ground state structure with the lowest energy [28,29]. Figure 5(a) demonstrates crystal structure models of pristine austenite with lattice constants of $a = b = c = 3.45$ Å [30]. Figure 5(b) shows the 2×2×2 austenite supercell doped with 3.125 at. % M atom (M = Cr, Ni, Mn), whose lattice constants of $a = b = c = 6.89$ Å.



(a) Austenite; (b) Austenite doped with M.

Figure 5. Crystal Structure Models.

3.3. The Convergence Test

Before geometric optimization, selecting the appropriate cutoff energy and K-point grid constants is crucial to ensuring the computational accuracy [31,32]. For all crystal structures, the cutoff energy and K-point parameters for geometry optimization were set to 600 eV and 6×6×6, respectively. Geometric optimization is performed using the GGA-PBESOL functional to relax lattice stresses and achieve atomic equilibrium. Table 2 presents the lattice constants and cell volumes of the undoped and the doped alloy. The calculated lattice constants show good consistency with the values in the relevant literature, thereby demonstrating the reliability of the present calculations. It can be observed that the crystal structure of the austenite undergoes varying degrees of distortion after doping, which exerts a significant influence on the stability of the austenite phase [17-19].

Table 2. Calculation values of lattice constants and cell volume of austenite and Cr/Ni/Mn-doped austenite.

Phases	**Method**	***a*=*b*=*c*/Å**	***V*/Å^3**
Austenite	GGA-PBE[30]	3.446	40.953
	Expt[33]	3.46	41.422
	Expt[34]	3.45	41.064
	GGA-PBE$_{\text{this work}}$	6.892	327.368
Austenite-Cr	GGA-PBE$_{\text{this work}}$	6.906	329.367
Austenite-Ni	GGA-PBE$_{\text{this work}}$	6.900	328.509
Austenite-Mn	GGA-PBE$_{\text{this work}}$	6.899	328.366

3.4. Stability of Doping System

To verify the prediction accuracy of the machine learning model. In this study, the system energy and solid solution energy were calculated for undoped austenite supercells and austenite supercells doped with Cr, Ni, and Mn, respectively, in order to assess the stability of the system before and after doping [35,36]. The solid solution energy is defined as follows:

$$E_{\text{Sol}} = E_{tot} - E(Fe) - E_{iso}(M) \quad (1)$$

Where E_{tot} is the total energy of the M-doped austenitic system, $E(Fe)$ is the energy of the austenite supercell after removing the doping atom, and $E_{iso}(M)$ is the energy of an isolated M atom.

Figure 6 shows the calculated values of the solid solution energy and system energy. Relevant studies show that a more negative solid solution energy indicates greater stability of the doping system [37]. The solid solution energies of austenite doped with Cr, Ni, and Mn are −11.79599, −11.79098, and −11.78801 eV·atom^{-1}, respectively. Meanwhile, the system energies of austenite doped with Cr, Ni, and Mn are −29609.59, −29147.37, and −28122.78 eV, respectively. This is consistent with the lattice distortion of doped austenite. The larger the

https://doi.org/10.64486/m.66.1.1

lattice constant of the doping element, the greater the lattice distortion, resulting in higher total energy and solid solution energy [38]. The above findings indicate that the larger the lattice constant of the doping element, the poorer the crystal stability of the doped system, thereby confirming the reliability of the results obtained from the machine learning model.

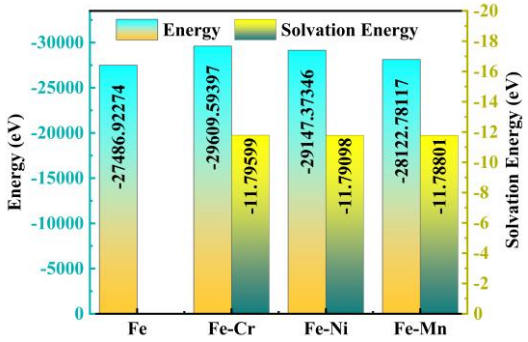


Figure 6. System energy and solid solution energy of austenite and Cr/Ni/Mn-doped austenite.

3.5. Elastic modulus

To further verify the mechanical stability of the doped system, the elastic constants including bulk modulus (K), Young's modulus (E), shear modulus (G), and Poisson's ratio (μ) are calculated systematically. The elastic matrix for austenite and Cr/Ni/Mn-doped austenite is calculated as follows:

$$C_{ij}(Gpa)=\begin{pmatrix} 432 & 182 & 182 & & & \\ 182 & 432 & 182 & & & \\ 182 & 182 & 431 & & & \\ & & & 250 & & \\ & & & & 250 & \\ & & & & & 250 \end{pmatrix} \quad C_{ij}(Gpa)=\begin{pmatrix} 425 & 224 & 224 & & & \\ 224 & 424 & 224 & & & \\ 224 & 224 & 424 & & & \\ & & & 239 & & \\ & & & & 239 & \\ & & & & & 239 \end{pmatrix}$$

$$C_{ij}(Gpa)=\begin{pmatrix} 363 & 169 & 169 & & & \\ 169 & 363 & 168 & & & \\ 169 & 169 & 363 & & & \\ & & & 241 & & \\ & & & & 241 & \\ & & & & & 241 \end{pmatrix} \quad C_{ij}(Gpa)=\begin{pmatrix} 418 & 187 & 187 & & & \\ 187 & 418 & 187 & & & \\ 187 & 187 & 418 & & & \\ & & & 251 & & \\ & & & & 251 & \\ & & & & & 251 \end{pmatrix}$$

Based on the above elastic constant matrix, the elastic constants of austenite supercells and austenite supercells doped with M (M= Cr, Ni, Mn) element were calculated, and the results are shown in Table 3. After doping, the bulk modulus, Young's modulus, and shear modulus of the material exhibit varying degrees of change. The bulk modulus of the material decreases to varying degrees as the lattice constant of the dopant element increases, indicating that the resistance to compression of the material is gradually decreasing. It is worth noting that the bulk modulus of the material increases when doped with elements of smaller lattice constant, indicating a negative correlation between lattice constant and system stability, which is consistent with the results of machine learning and correlation analysis. Moreover, the Poisson's ratio of the doped system exhibits varying degrees of change, indicating that the ductility of the material has also changed.

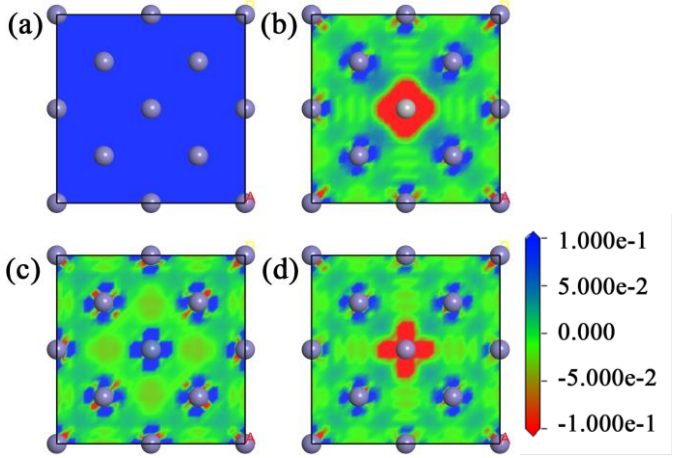
https://doi.org/10.64486/m.66.1.1

Table 3. The elastic constants of austenite and Cr/Ni/Mn-doped austenite.

Phases	Bulk modulus/GPa	Young's modulus/GPa	Shear modulus/GPa	Poisson's ratio
Austenite	265.33	479.52	200.00	0.199
Austenite-Cr	291.00	455.52	183.6	0.239
Austenite-Ni	233.67	436.07	183.4	0.189
Austenite-Mn	264.00	473.96	196.8	0.201

3.6. Charge density differences

To further clarify the stability of the doped systems and verify the predictions of machine learning model, the charge density difference of austenite doped with M (M = Cr, Ni, Mn) are investigated, as presented in Figure 7. The charge density difference in the range of –0.1 to 0.1 e·Å^{-3} is represented by a color map with red indicating charge depletion and blue indicating charge accumulation. As shown in Figure 7(a), the charge density difference map of austenite exhibits a uniform charge distribution. However, the charge density difference of the doped system has changed significantly. Relevant studies show that the appearance of extensive electron clouds indicates an enhancement of the interatomic interaction [39]. For the Cr-doped austenite system, the Cr atom loses a significant number of electrons, while its adjacent Fe atoms gain electrons, as shown in Figure 7(b). Moreover, an obvious electron cloud distribution is observed between the Cr and Fe atoms, indicating a strong interaction between them. Furthermore, a slight charge accumulation is also observed between distant iron atoms, indicating that the Fe-Fe interaction is enhanced as well. A comparative analysis of the charge density differences among the three doped systems reveals that the system doped with the element having the smallest lattice constant exhibits the most pronounced electron cloud distribution, thereby validating the results of the machine learning predictions. In summary, the charge density differences further confirm that each M (M = Cr, Ni, Mn) atom promotes increased electron sharing with its neighboring Fe atoms, collectively contributing to enhanced structural stability.



(a) Austenite; (b) Austenite -Cr; (c) Austenite -Ni; (d) Austenite -Mn.

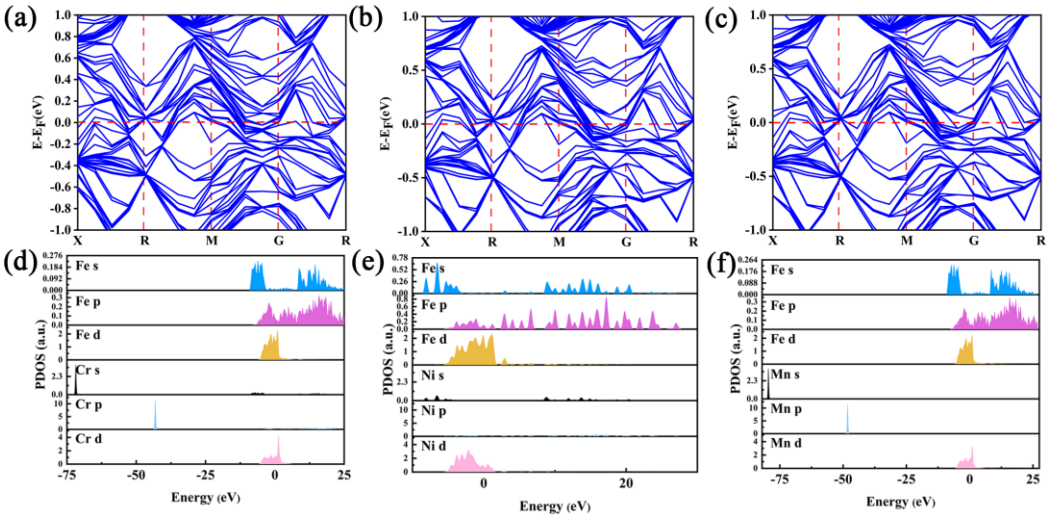
Figure 7. Charge density differences.

3.7. Structural properties

Calculations of electronic properties including band structures and partial density of states (PDOS), which provide critical insights into electron behavior and energy states, are essential for understanding material properties [40,41]. Consequently, calculating the above-mentioned material properties helps to further verify the predicted results. The band structures of austenite doped with Cr, Ni, and Mn are shown in Figure 8(a-c). The results show that the valence band crosses the Fermi level (the dashed line at 0 eV) and enters the conduction band region, indicating that a large number of free electrons exist near the Fermi level. Figure 8(d-f) presents the PDOS results for the doped system. The results show that the M-d orbitals hybridize significantly with the Fe-p and Fe-d orbitals near the Fermi level, indicating a strong covalent bonding interaction between the dopant

https://doi.org/10.64486/m.66.1.1

and iron. Furthermore, it is observed that the hybridization near the Fermi level gradually weakens as the lattice constant of the dopant element increases. The above findings confirm the machine learning prediction that the larger the lattice constant of the added M element, the weaker the stability of the austenite crystal, and this variation is attributed to the changes in atomic interactions.



(a) Austenite -Cr; (b) Austenite -Ni; (c) Austenite -Mn; density of states: (d) Austenite -Cr; (e) Austenite -Ni; (f) Austenite -Mn.

Figure 8. Band structure.

4. Conclusions

This study systematically investigates the effect of doped elements on the stability of the austenitic system by integrating machine learning with first-principles calculations, providing a novel approach for developing steel materials. The main conclusions are summarized as follows:

(1) Pearson correlation coefficient analysis shows that the lattice constant is the most influential feature for the stability of doped systems, with an R value of 0.308. Furthermore, among the three machine learning models evaluated, the RF model achieves the highest prediction accuracy, with an R^2 of 0.748.

(2) Analysis of the system energy and solid solution energy of the doped systems shows that as the lattice constant of the dopant element increases, the system energy and solid solution energy gradually rise. Among them, the Cr-doped austenitic system exhibits the lowest system energy and solution energy, with values of -29609.59 eV and -11.795 eV, respectively. The order of system stability corresponding to the dopant elements is Cr > Ni > Mn, which is consistent with the prediction results of the machine learning model.

(3) Analysis of the charge density differences indicates that the electron cloud density between atoms in the doped systems is enhanced, suggesting stronger interatomic interactions compared with the original system. Band structure and density of states analyses reveal significant free electrons near the Fermi level, as well as orbital hybridization between Fe and dopant atoms, which suggests covalent bonding interactions that enhance structural stability. Furthermore, as the lattice constant increases, the aforementioned interactions all exhibit a decreasing trend.

(4) This study integrates machine learning with first-principles calculations, offering a robust framework for accelerating the discovery of novel austenitic steels.

https://doi.org/10.64486/m.66.1.1

Acknowledgments: This work was supported by the Innovation Capability Enhancement Plan Project of Hebei Province (No. 24461002D), National Natural Science Foundation of China (No. 52304351), Project of Yanzhao Iron and Steel Laboratory (NO. YZISL2024001), Science and the Central Guidance on Local Science and Technology Development Fund of Hebei Province (No.226Z1006G).

References

[1] M. T. Chen, Z. C. Gong, T. Y. Zhang, et al., "Mechanical behavior of austenitic stainless steels produced by wire arc additive manufacturing," *Thin-Walled Struct*., vol. 196, Art. no. 111455, 2024, https://doi.org/10.1016/j.tws.2023.111455

[2] H. J. Mei, L. Y. Cheng, L. Chen, et al., "Effect of alloying elements on zinc-induced liquid metal embrittlement in steels: A first-principles study," *Comput. Mater. Sci*., vol. 242, Art. no. 113104, 2024, https://doi.org/10.1016/j.commatsci.2024.113104

[3] Y. F. Su, S. Z. Zhang, S. X. Jiao, et al., "Nitrogen enhances microstructural thermal stability of Si-modified Fe-Cr-Ni austenitic stainless steel," *J. Mater. Sci. Technol*., vol. 226, pp. 270-289, 2025, https://doi.org/10.1016/j.jmst.2024.11.027

[4] M. Morales, M. Rezayat, G. Fargas, et al., "Mitigating the corrosion of AISI 301LN steel in molten carbonate salts by doping with alumina nanoparticles for thermal energy storage applications," *Sol. Energy Mater. Sol. Cells*, vol. 270, Art. no. 112805, 2024, https://doi.org/10.1016/j.solmat.2024.112805

[5] S. M. Yang, Y. C. Chen, Y. T. Pan, et al., "Effect of silver on microstructure and antibacterial property of 2205 duplex stainless steel," *Mat. Sci. Eng. C Mater*., vol. 63, pp. 376-383, 2016, https://doi.org/10.1016/j.msec.2016.03.014

[6] L. F. Jiang, Y. N. Hu, Y. X. Liu, et al., "Physics-informed machine learning for low-cycle fatigue life prediction of 316 stainless steels," *Int. J. Fatigue*, vol. 182, Art. no. 108187, 2024, https://doi.org/10.1016/j.ijfatigue.2024.108187

[7] J. Zhang, S. K. Zhang, X. Z. Xu, et al., "First-principles calculation of Fe-Mo-Cr alloys based on special quasi-random structure," *Mater. Today Commun*., vol. 35, Art. no. 106179, 2023, https://doi.org/10.1016/j.mtcomm.2023.106179

[8] Elements: The Periodic Table, 2023. [Online]. Available: https://periodictable.chemicalaid.com/ [Accessed: Jun. 27, 2025]

[9] Y. Zhang, C. Ling, "A strategy to apply machine learning to small datasets in materials science," *npj Comput. Mater*., vol. 4, no. 1, Art. no. 25, 2018, https://doi.org/10.1038/s41524-018-0081-z

[10] K. Min, B. Choi, K. Park, et al., "Machine learning assisted optimization of electrochemical properties for Ni-rich cathode materials," *Sci. Rep*., vol. 8, Art. no. 15778, 2018, https://doi.org/10.1038/s41598-018-34201-4

[11] M. Raissi, P. Perdikaris, G. E. Karniadakis, "Machine learning of linear differential equations using Gaussian processes," *J. Comput. Phys*., vol. 348, pp. 683-693, 2017, https://doi.org/10.1016/j.jcp.2017.07.050

[12] D. Singh, B. Singh, "Feature wise normalization: An effective way of normalizing data," *Pattern Recognit*., vol. 122, Art. no. 108307, 2022, https://doi.org/10.1016/j.patcog.2021.108307

[13] K. Choudhary, B. DeCost, C. Chen, et al., "Recent advances and applications of deep learning methods in materials science," *npj Comput. Mater*., vol. 8, no. 1, Art. no. 59, 2022, https://doi.org/10.1038/s41524-022-00734-6

[14] B. Kalaiselvi, M. Thangamani, "An efficient Pearson correlation based improved random forest classification for protein structure prediction techniques," *Measurement*, vol. 162, Art. no. 107885, 2020, https://doi.org/10.1016/j.measurement.2020.107885

[15] S. Z. Zhang, H. Y. Zhu, Y. F. Su, et al., "High temperature stress relaxation behavior of high Si, Mo-doped austenitic stainless steels," *Mat. Sci. Eng. A Struct*., vol. 916, Art. no. 147330, 2024, https://doi.org/10.1016/j.msea.2024.147330

[16] B. S. Mi, T. J. Chen, X. Ma, et al., "Improving corrosion and conductivity of Ti-doped amorphous carbon film coated on 316L stainless steel bipolar plates by nitrogen atoms doping," *Vacuum*, vol. 228, Art. no. 113496, 2024, https://doi.org/10.1016/j.vacuum.2024.113496

[17] X. X. Wu, Z. J. Sun, W. Q. Guo, et al., "Microstructural origin of the simultaneous enhancements in strength and ductility of a nitrogen-doped high-entropy alloy," *Acta Mater*., vol. 304, Art. no. 121753, 2026, https://doi.org/10.1016/j.actamat.2025.121753

[18] C. H. Ng, M. J. Bermingham, L. Yuan, et al., "Towards β-fleck defect free additively manufactured titanium alloys by promoting the columnar to equiaxed transition and grain refinement," *Acta Mater*., vol. 224, Art. no. 117511, 2022, https://doi.org/10.1016/j.actamat.2021.117511

[19] D. Wu, Y. Yang, X. F. Zhang, et al., "First-principles study on effect of alloying elements on heterogeneous nucleation of reverse austenite on Cu precipitation," *Acta Phys. Sin*., vol. 71, no. 8, pp. 275-283, 2022, https://doi.org/10.7498/aps.71.20212144

[20] X. P. Li, Y. D. Wang, R. B. Ruiz, et al., "A Survey on Sparse Learning Models for Feature Selection, IEEE Trans. *Cybern*.," vol. 52, no. 3, pp. 1642-1660, 2020, https://doi.org/10.1109/TCYB.2020.2982445

[21] M. Kim, S. Kang, H. G. Park, et al., "Maximizing the energy density and stability of Ni-rich layered cathode materials with multivalent dopants via machine learning," *Chem. Eng*. J., vol. 452, Art. no. 139254, 2023, https://doi.org/10.1016/j.cej.2022.139254

[22] M. Sabzekar, S. Hasheminejad, "Robust regression using support vector regressions," *Chaos Soliton. Fract*., vol. 144, Art. no. 110738, 2021, https://doi.org/10.1016/j.chaos.2021.110738

[23] A. V. Ponce-Bobadilla, V. Schmitt, C. S. Maier, et al., "Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development," *Clin. Transl. Sci*., vol. 17, Art. no. e70056, 2024, https://doi.org/10.1111/cts.70056

[24] V. E. Ol'shanetskii, G. V. Snezhnoi, V. N. Sazhnev, "Structural and Magnetic Stability of Austenite in Chromium-Nickel and Manganese Steels with Cold Deformation," *Met. Sci. Heat Treat*., vol. 58, no. 5-6, pp. 261-269, 2016, https://doi.org/10.1007/s11041-016-0009-5

[25] J. Hullman, P. Resnick, E. Adar, "Hypothetical Outcome Plots Outperform Error Bars and Violin Plots for Inferences about Reliability of Variable Ordering," *PLoS ONE*, vol. 10, Art. no. e0142444, 2015, https://doi.org/10.1371/journal.pone.0142444

[26] Y. Kobayashi, T. Akasaka, "Hexagonal BN epitaxial growth on (0 0 0 1) sapphire substrate by MOVPE," *J. Cryst. Growth*, vol. 310, no. 23, pp. 5044-5047, 2008, https://doi.org/10.1016/j.jcrysgro.2008.07.010

[27] P. Kroll, M. Milko, "Theoretical Investigation of the Solid State Reaction of Silicon Nitride and Silicon Dioxide forming Silicon Oxynitride (Si2N2O) under Pressure," *Z. Anorg. Allg. Chem*., vol. 629, no. 10, pp. 1737-1750, 2003, https://doi.org/10.1002/zaac.200300122

[28] V. Milman, K. Refson, S. J. Clark, et al., "Electron and vibrational spectroscopies using DFT, plane waves and pseudopotentials: CASTEP implementation," *J. Mol. Struct. Theochem*, vol. 954, no. 1-3, pp. 22-35, 2010, https://doi.org/10.1016/j.theochem.2009.12.040

[29] S. G. Christopoulos, K. A. Papadopoulou, A. Konios, et al., "DIMS: A tool for setting up defects and impurities CASTEP calculations," *Comput. Mater. Sci*., vol. 202, Art. no. 110976, 2022, https://doi.org/10.1016/j.commatsci.2021.110976

[30] D. L. Zhang, J. Wang, Y. Kong, et al., "First-principles investigation on stability and electronic structure of Sc-doped θ′/Al interface in Al–Cu alloys," *Trans. Nonferrous Met. Soc*. China, vol. 31, no. 11, pp. 3342-3355, 2021, https://doi.org/10.1016/S1003-6326(21)65733-3

[31] D. Adekoya, S. S. Qian, X. X. Gu, et al., "DFT-Guided Design and Fabrication of Carbon-Nitride-Based Materials for Energy Storage Devices: A Review," *Nano-Micro Lett*., vol. 13, no. 13, pp. 1-44, 2021, https://doi.org/10.1007/s40820-020-00522-1

[32] M. Bursch, J. M. Mewes, A. Hansen, et al., "Best-Practice DFT Protocols for Basic Molecular Computational Chemistry," *Angew. Chem*., vol. 134, no. 42, Art. no. e202205735, 2022, https://doi.org/10.1002/ange.202205735

[33] L. L. Liu, R. Wang, X. Z. Wu, et al., "Temperature effects on the generalized planar fault energies and twinnabilities of Al, Ni and Cu: First principles calculations," *Comput. Mater*. Sci., vol. 88, pp. 124-130, 2014, https://doi.org/10.1016/j.commatsci.2014.03.005

[34] G. J. Chen, S. D. Guo, H. H. Zhang, et al., "The effects of active elements on adhesion strength of SiC/Cu interface in SiC reinforced Cu-based composite: A first-principles investigation", Mater. *Today Commun*., vol. 31, Art. no. 103233, 2022, https://doi.org/10.1016/j.mtcomm.2022.103233

[35] Y. You, J. H. Yan, M. F. Yan, et al., "La interactions with C and N in bcc Fe from first principles," *J. Alloys Compd*., vol. 688, pp. 261-269, 2016, https://doi.org/10.1016/j.jallcom.2016.07.015

[36] Y. H. Mo, Z. Q. Tang, L. Zhang, et al., "Data-driven discovery of vanadium-based anode materials for lithium-ion batteries," *J. Energy Storage*, vol. 118, Art. no. 116290, 2025, https://doi.org/10.1016/j.est.2025.116290

[37] Y. Yang, X. Z. Dai, X. R. Yang, et al., "First-principles analysis on the role of rare-earth doping in affecting nitrogen adsorption and diffusion at Fe surface towards clarified catalytic diffusion mechanism in nitriding," *Acta Mater*., vol. 196, pp. 347-354, 2020, https://doi.org/10.1016/j.actamat.2020.06.020

[38] D. Fan, C. Liu, F. Yin, et al., "Solid solution strengthening mechanism and interstitial diffusion behavior of rare earth element lanthanum in austenite using first-principles calculations", *Theor. Chem. Acc*., vol. 139, pp. 3, 2020, https://doi.org/10.1007/s00214-019-2513-3

[39] N. Kai, L. Fengde, L. Xingran, et al., "First principle study on the effect of Cr, Mn and Ni on electronic structure and mechanical properties of high nitrogen steel," *Appl. Phys. A*, vol. 129, pp. 109, 2023, https://doi.org/10.1007/s00339-022-06167-8

[40] B. O. Mnisi, "Ab initio study of phase stability and optical properties of TiN and VN nitrides in different phases," *Solid State Commun*., vol. 377, Art. no. 115390, 2024, https://doi.org/10.1016/j.ssc.2023.115390

[41] H. Z. Liu, S. K. Zhang, J. Zhang, et al., "Properties of typical non-metallic inclusions in steel: First-principles calculations," *Mater. Today Commun*., vol. 34, Art. no. 105118, 2023, https://doi.org/10.1016/j.mtcomm.2022.105118