

Machine learning accelerated search for superconductors in B-C-N based compounds and $R_3\text{Ni}_2\text{O}_7$ -type nickelates

Xiaoying Li,^{1,2} Wenqian Tu,^{1,2} Run Lv,^{1,2} Li'e Liu,^{1,2} Dingfu Shao,¹ Yuping Sun,^{1,3,4} and Wenjian Lu^{1,*}
¹Key Laboratory of Materials Physics, *Institute of Solid State Physics, HFIPS, Chinese Academy of Sciences, Hefei 230031, China*
²Science Island Branch of Graduate School, *University of Science and Technology of China, Hefei 230026, China*
³Anhui Province Key Laboratory of Low-Energy Quantum Materials and Devices, *High Magnetic Field Laboratory, HFIPS, Chinese Academy of Sciences, Hefei 230031, China*
⁴*Collaborative Innovation Center of Microstructures, Nanjing University, Nanjing 210093, China*



(Received 7 September 2025; revised 5 January 2026; accepted 3 February 2026; published 24 February 2026)

Superconductor research has traditionally depended on experiments and theoretical approaches. However, the rapid advancement of data-driven methods and machine learning (ML) has opened avenues for accelerating superconductor discovery. Here, we integrate ML with density functional theory calculations to efficiently screen conventional B-C-N based superconductors and identify potential high- T_C candidates among $R_3\text{Ni}_2\text{O}_7$ -type bilayer nickelates. We identify 13 binary and ternary B-C-N based superconductors with $T_C \geq 10$ K, including 3 with $T_C \geq 25$ K, two structural forms of B_2CN ($T_C = 47.6$ and 43.7 K) and TiNbN_2 ($T_C = 25.3$ K). These B-C-N based compounds share a common feature of strong σ bonds, which is key to achieving relatively high T_C . Moreover, we propose $\text{Tb}_3\text{Ni}_2\text{O}_7$ ($T_C = 61.6$ K) and $\text{Ac}_3\text{Ni}_2\text{O}_7$ ($T_C = 70.3$ K) as potential high- T_C nickelate superconductors under high pressure. Their electronic structures closely resemble those of $\text{La}_3\text{Ni}_2\text{O}_7$, especially in the hole-type band dominated by Ni $3d_z^2$ orbital character. We also analyze feature importance in the ML results for both conventional and high- T_C superconductors. These results advance the search for new superconductors and enhance the fundamental understanding of superconducting mechanisms.

DOI: [10.1103/wn98-9yc7](https://doi.org/10.1103/wn98-9yc7)

I. INTRODUCTION

Since the discovery of superconductivity in mercury by Onnes in 1911 [1], the pursuit of materials with high critical temperatures T_C has been a hot topic in condensed matter physics. Conventional superconductivity is explained by the BCS theory [2], which proposes that light atomic masses can lead to a higher Debye frequency and thus enhance T_C . A prominent example is MgB_2 , which holds the record T_C of 39 K among conventional superconductors [3]. In recent decades, T_C values have risen dramatically with the discovery of cuprate, iron-based, and nickelate superconductors, many of which exceed the boiling temperature of liquid nitrogen (77 K). Notably, the recent observation of superconductivity at 80 K in $\text{La}_3\text{Ni}_2\text{O}_7$ under pressure [4] has attracted broad interest. While the square Ni-O layers in $\text{La}_3\text{Ni}_2\text{O}_7$ resemble the Cu-O planes in cuprates, a key difference lies in the bonding of nickel atoms between adjacent layers via apical oxygen [5]. The intricate electronic correlations and layered crystal structures in these high- T_C systems pose significant challenges for understanding superconducting mechanisms and discovering new superconductors.

In recent years, materials science has moved into its fourth paradigm: “data-based materials science” [6], opening new avenues for accelerating materials discovery. Traditional approaches, which rely heavily on experimental trial and error and theoretical calculations [7], are often slow, computation-

ally expensive, and inefficient for high-throughput screening. In contrast, data-driven and machine learning (ML) methods offer powerful tools for predicting and exploring superconducting materials. For instance, data-guided design led to the identification of MgB_2 -like superconductors that were later confirmed experimentally [8]. Several studies have used ML regression methods to predict T_C of new compounds [9,10], and more methods, such as convolutional neural networks, have been employed to improve the prediction accuracy, including for unconventional transition metal oxides [11].

Nevertheless, many such predictions remain purely numerical estimates of T_C , lacking mechanistic insight or experimental validation. To address these limitations, researchers have developed integrated frameworks combining ML with density functional theory (DFT) methods to enhance the reliability of predictions for conventional superconductors. For example, Hoffmann *et al.* conducted a systematic investigation of antiperovskites [12] and full-Heusler compounds [13], identifying multiple correlated superconductors. Similar approaches have been applied to high-pressure hydride systems [14–17]. These investigations demonstrate the potential of ML not only for accelerating discovery but also for mapping uncharted regions of the chemical space.

Motivated by the BCS theory and guided by statistical trends in conventional superconductors (see Fig. S1 of the Supplemental Material [18]) and the recent discovery of superconductivity in $\text{La}_3\text{Ni}_2\text{O}_7$, we employ an integrated ML and DFT framework to investigate the superconducting potential of binary and ternary B-C-N based compounds and bilayer nickelates $R_3\text{Ni}_2\text{O}_7$ (R = rare-earth elements). We identified

*Contact author: wjlu@issp.ac.cn

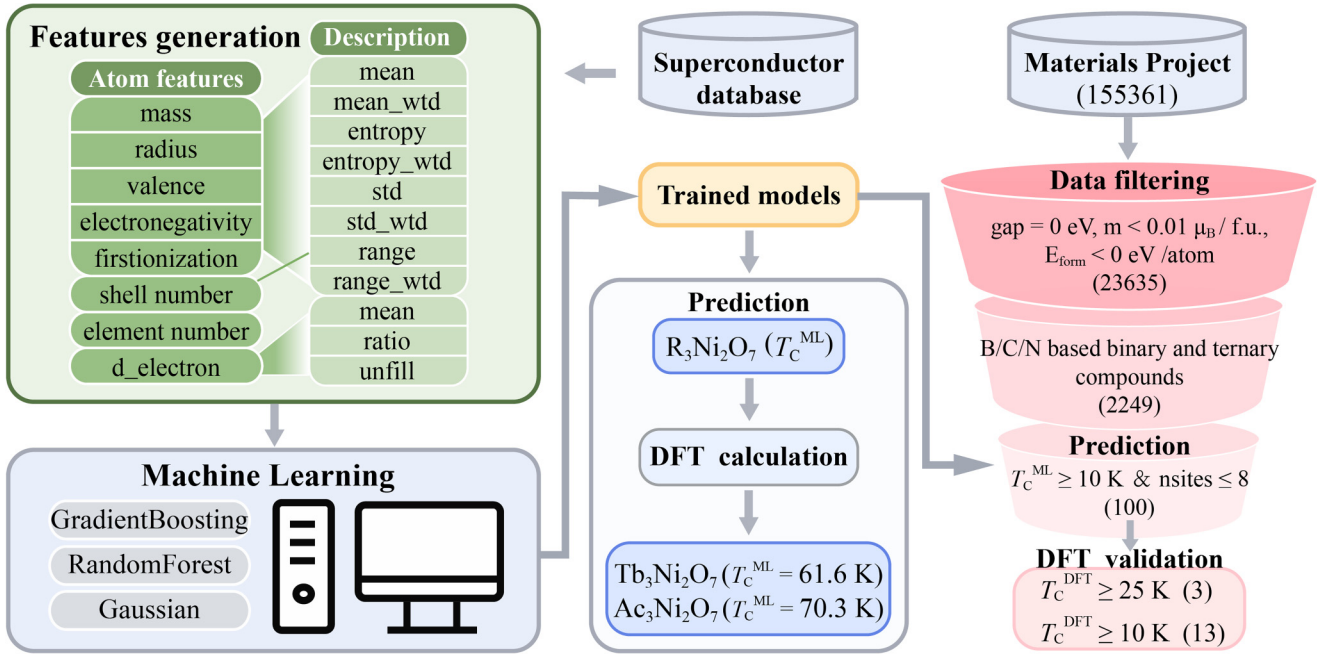


FIG. 1. Workflow for accelerating the discovery of potential superconductors. The workflow involves generating features for experimentally discovered superconductors as learning data, training ML models, filtering candidate materials based on ML-predicted T_C , and validating predictions using DFT calculations.

13 binary and ternary B-C-N based conventional superconductors with $T_C > 10$ K, including two structural forms of B_2CN ($T_C = 47.6$ K and 43.7 K) and $TiNbN_2$ ($T_C = 25.3$ K). These materials exhibit strong σ -bonding characteristics, similar to those driving superconductivity in MgB_2 . Our ML model also predicts $Tb_3Ni_2O_7$ and $Ac_3Ni_2O_7$ are potential superconducting candidates, with estimated T_C values of 61.6 and 70.3 K at 30 GPa, respectively. Electronic structure analysis reveals that the hole-type bands contributed by the Ni $3d_z^2$ orbital in $Tb_3Ni_2O_7$ and $Ac_3Ni_2O_7$ resemble those in $La_3Ni_2O_7$, with a similar hole concentration ($n_{hole} \sim 0.135$). By interpreting key features identified by the ML model, we gain deeper insight into the mechanisms underlying both conventional and high- T_C superconductivity. This work not only shows how integrated ML-DFT frameworks can simultaneously achieve reliable predictions of superconductors but also deepens our understanding of their fundamental mechanisms, thereby bridging the gap between ML screening and physical insight.

II. METHODOLOGY

A. Workflow

Figure 1 presents our workflow, which integrates ML methods and DFT calculations, enabling the efficient identification of superconducting candidates. The specific details of this workflow will be explained in the subsequent sections.

B. Learning data preparation

High-quality learning data are essential for developing robust ML models. In this study, we utilize the SuperCon database [19], a rigorously curated resource maintained by the National Institute for Materials Science of Japan. This

database systematically archives experimentally validated superconductors and the associated literature. Our learning dataset includes 5870 conventional superconductors and 3940 cuprate high-temperature superconductors, with only two key attributes retained for each compound: the chemical formula and experimental T_C . The dataset exhibits a broad T_C range: 0–41.4 K for conventional superconductors and 0–143.0 K for cuprates. Introducing diverse superconducting materials in training improves the accuracy and generalization of the ML models.

The foundation of feature generation lies in selecting physically interpretable elemental properties. We first created eight fundamental atomic attributes for each element: the number of elements, atomic mass, atomic radius, valence, electronegativity, first ionization energy, number of electron shells, and number of d electrons. These properties have a critical influence on bonding dynamics and electronic interactions within crystals. For instance, the atomic mass and radius affect lattice vibrations (phonons), while valence electrons, electronegativity, and first ionization energy modulate charge transfer. To connect atomic-scale properties with macroscopic behavior, we followed a previous study [9] and applied a series of mathematical operations to these properties, as detailed in Table I. Through systematic combinations, each chemical formula possesses 45 distinct features. The transformed feature space offers two key advantages for ML models: It captures interactions among elements in complex bonding networks, reflecting real material behavior, and provides richer, more comprehensive material-specific information for enhanced learning.

C. DFT calculation details

Electronic property calculations based on DFT were implemented via the Vienna *Ab initio* Simulation Package

TABLE I. Summary of the procedure for feature generation from the material's chemical formula. The last column serves as an example: Operations are performed on the features based on the atomic mass extracted for MgB_2 . Mg and B atomic mass values are $t_1 = 24.305$ and $t_2 = 10.811$, respectively. Here, $p_1 = \frac{1}{1+2} = \frac{1}{3}$, $p_2 = \frac{2}{1+2} = \frac{2}{3}$, $w_1 = \frac{t_1}{t_1+t_2} \approx 0.692$, $w_2 = \frac{t_2}{t_1+t_2} \approx 0.308$, $A = \frac{p_1 w_1}{p_1 w_1 + p_2 w_2} \approx 0.529$, and $B = \frac{p_2 w_2}{p_1 w_1 + p_2 w_2} \approx 0.471$.

Description	Symbol	Formula	Sample value
Mean	mean	$\mu = (t_1 + t_2)/2$	17.558
Weighted mean	mean_wtd	$v = (p_1 t_1 + p_2 t_2)$	15.309
Entropy	entropy	$-w_1 \ln(w_1) - w_2 \ln(w_2)$	0.617
Weighted entropy	entropy_wtd	$-A \ln(A) - B \ln(B)$	0.691
Standard deviation	std	$[1/2((t_1 - \mu)^2 + (t_2 - \mu)^2)]^{1/2}$	6.747
Weighted standard deviation	std_wtd	$[p_1(t_1 - v)^2 + p_2(t_2 - v)^2]^{1/2}$	6.361
Range	range	$t_1 - t_2$ ($t_1 > t_2$)	13.494
Weighted range	range_wtd	$p_1 t_1 - p_2 t_2$	0.894

(vasp) [20,21]. The generalized gradient approximation of the Perdew-Burke-Ernzerhof form was employed as the exchange-correlation functional. A fine k -point mesh with a resolution of $0.02 \text{ \AA}^{-1}$ was used for the self-consistent field (SCF) calculations, while a denser mesh with a resolution of $0.005 \text{ \AA}^{-1}$ was employed for Fermi surface calculations. The projector augmented-wave method with a 600 eV plane-wave kinetic cutoff energy was used. The phonon calculations were performed by using the quantum espresso package [22] with norm-conserving pseudopotentials. The cutoff energies of the wave functions and charge density were set as 60 and 600 Ry, respectively. To accurately estimate T_C , the electron-phonon coupling (EPC) properties were calculated using the epw code [23], which employs the Wannier-Fourier interpolation method with denser k/q -point meshes. The electron-phonon matrix was interpolated onto a $30 \times 30 \times 30$ grid for both k and q points. To better describe the strong correlation of Ni $3d$ electrons in bilayer nickelates $R_3\text{Ni}_2\text{O}_7$, we used the DFT+U method, with Hubbard $U = 3$ eV and Hund's parameter $J = 0.4$ eV, which was also adopted in a previous study [24]. The $4f$ electrons of R atoms were considered as core electrons. For conventional superconductors, we used the Allen-Dynes-modified McMillan equation [25], $T_C = \frac{\omega_{\log}}{1.2} \exp[\frac{-1.04(1+\lambda)}{\lambda(1-0.62\mu^*)-\mu^*}]$, where EPC strength $\lambda = \sum_{qv} \lambda_{qv} \int \frac{\alpha^2 F(\omega)}{\omega} d\omega$ and the Eliashberg spectral function $\alpha^2 F(\omega) = \frac{1}{2\pi N(E_F)} \sum_{qv} \delta(\omega - \omega_{qv}) \frac{\gamma_{qv}}{\hbar \omega_{qv}}$, to quantitatively evaluate T_C .

III. RESULTS AND DISCUSSION

A. Model training

Using the prepared superconductor datasets with 45 distinct features, we trained three ML models—gradient boosting, random forest, and Gaussian regression models—to evaluate their performance in estimating T_C of superconductors. Each model has unique characteristics and strengths. We rigorously compared their performance using three critical parameters: mean absolute error (MAE), root-mean-square error (RMSE), and the coefficient of determination R^2 . The corresponding calculation formulas are shown in Eqs. (1)–(3). This

comprehensive evaluation enables objective identification of the most appropriate model.

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|, \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}, \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2}. \quad (3)$$

The dataset, comprising 5870 conventional superconducting materials, was divided into a training set (90%) and a testing set (10%) following the standard protocol. Three ML models were trained and evaluated on the dataset. The training results show that the gradient boosting model outperformed the others, achieving $\text{MAE} = 1.128$, $\text{RMSE} = 2.139$, and $R^2 = 0.919$ on the testing set [Fig. 2(a)]. The evaluation results, as shown in Fig. 2(b), reveal the model's predictive capability and generalization stability, with $R^2(\text{train}) = 0.99$ and $R^2(\text{test}) = 0.92$.

To better investigate nickelates, we constructed a hybrid dataset integrating 3940 cuprates from the SuperCon database and 126 nickelate materials (containing different reported T_C values for LaNiO_2 , $\text{La}_3\text{Ni}_2\text{O}_7$, and $\text{La}_4\text{Ni}_3\text{O}_{10}$ -type materials at various pressures) from the literature. A pressure variable was added as a new feature alongside the 45 existing features. Among the three models employed in training, the gradient boosting model also demonstrated the best generalization performance, with $R^2(\text{train}) = 0.96$ and $R^2(\text{test}) = 0.83$ [Fig. 2(c)]. Consequently, the model provided robust support for advancing in-depth investigation of conventional and high- T_C superconducting materials.

To further validate the prediction accuracy of our model, we applied the trained model to predict T_C of several recently reported experimental superconductors (not included in the learning dataset). As listed in Table II, the ML-predicted T_C (T_C^{ML}) values show good agreement with the experimentally observed T_C ($T_C^{\text{exp.}}$) values. Such consistency indicates that our trained model can deliver accurate predictions, even for materials not included in the training dataset.

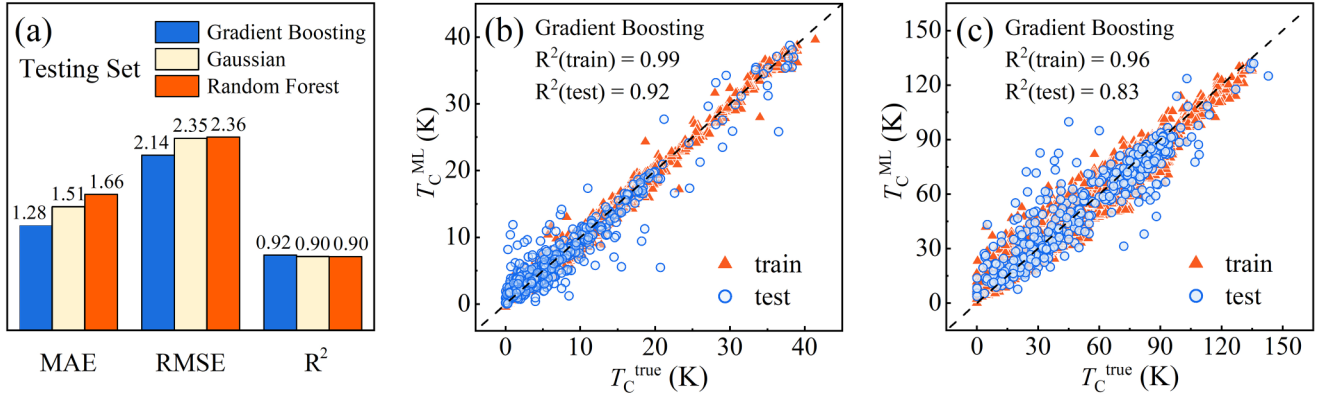


FIG. 2. Evaluation of ML models. (a) Blue, yellow, and orange denote gradient boosting, Gaussian, and random forest regression models for training conventional superconductors, respectively. (b) and (c) Benchmarking of gradient boosting regression models for predicting T_C of conventional superconductors and high- T_C superconductors, respectively. The orange triangles and blue circles represent the training set and testing set data, respectively.

B. Feature importance discussion

Feature importance analysis in Fig. 3(a) identifies that the weighted mean mass (mass_mean_wtd), weighted mean valence (valence_mean_wtd), and weighted mean first ionization energy (firstionization_mean_wtd) are the three most critical variables for conventional superconductors, consistent with established superconductivity theory and prior experimental observations. As is well-known, conventional superconductivity relies on atomic mass (reflected by mass_mean_wtd) to modulate lattice vibration frequencies, thereby regulating EPC strength and, ultimately, determining the T_C . As proposed by Reynolds *et al.*, the empirical relationship demonstrates that $T_C \propto 1/\sqrt{m}$ [37]. As shown in Fig. 3(b), the weighted average attribute of mass_mean_wtd effectively captures the combined influence of different atoms in the compound on lattice vibrations.

The second key feature, valence_mean_wtd, corresponds to the valence electron concentration (VEC) [38]: $\text{VEC} = (me_C + ne_A) / (m + n)$, where m and n represent the numbers of cations and anions and e_C and e_A represent the numbers of the cation and anion valence electrons. VEC

directly influences the density of charge carriers available for pairing and thus the magnitude of the superconducting energy gap. This relationship echoes the empirical finding first discovered by Matthias more than 70 years ago [39]. It revealed that for early discovered superconductors, two maxima in T_C were found for VECs near 5 and 7. Later, MgB_2 -related superconductors were found and associated with a T_C peak for a VEC near 3. Our analysis of valence_mean_wtd in Fig. 3(c) shows that there are optimal VEC values for distinct families: approximately 3 for MgB_2 -related compounds, 4 for doped C_{60} , and 5 for bismuthates, while A15 family materials exhibit double peaks in their T_C at VECs of ~ 4.5 and 6.5 .

The first ionization energy, quantifying the energy required to remove the first electron, can significantly influence a material's electronic configuration, its conductivity, and, ultimately, its superconducting properties. Although there was no significant linear correlation between firstionization_mean_wtd and T_C overall, specific values are notable in particular systems: approximately 8 for MgB_2 -related compounds, 11 for doped C_{60} , 10.5 for bismuthates, and 7 for A15 family materials.

We further analyze the features for high- T_C superconductors and show the results in Fig. 4. The weighted standard deviation of valence electrons (valence_std_wtd) has the highest weight, indicating that the dispersion of valence electron distributions within compounds plays a dominant role in regulating T_C . Other important features include the weighted standard deviation of the atomic radius (radius_std_wtd) and the number of unfilled d -orbital electrons (d_{unfill}). Figures 4(b)–4(d) illustrate how the three key features correlate with T_C across diverse material systems.

Figure 4(b) reveals a nonmonotonic relationship of T_C vs valence_std_wtd: T_C initially increases with rising valence_std_wtd, peaks at valence_std_wtd ≈ 3.25 , and then declines. Despite the evident overall trend, significant scatter persists among the data points. This variability indicates that although valence_std_wtd is a primary driver of T_C variations, the chemical environment—including factors such as lattice distortions and interatomic interactions—may further refine the relationship between valence electron dispersion and T_C .

TABLE II. Comparisons of ML-predicted T_C and experimentally observed T_C of recently discovered superconductors (not included in learning data).

Formula	T_C^{ML} (K)	$T_C^{\text{exp.}}$ (K)	Ref.
V_2AlN	9.3	15.9	[26]
$\text{Ta}_2\text{V}_{3.1}\text{Si}_{0.9}$	6.9	7.5	[27]
$\text{Nb}_3\text{P}_2\text{Se}_2$	6.9	6.8	[28]
La_2IRu_2	6.2	4.8	[29]
$\text{W}_3\text{Al}_2\text{C}$	6.2	7.6	[30]
$\text{ZrP}_{1.27}\text{Se}_{0.73}$	6.0	7.1	[31]
Nb_2PSe_2	5.8	6.9	[28]
InNbS_2	5.4	6.0	[32]
TlBi_2	5.4	6.2	[33]
NbReSi	4.7	6.5	[34]
Mo_3ReRuC	4.5	7.7	[35]
ThRu_3Si_2	4.1	3.8	[36]

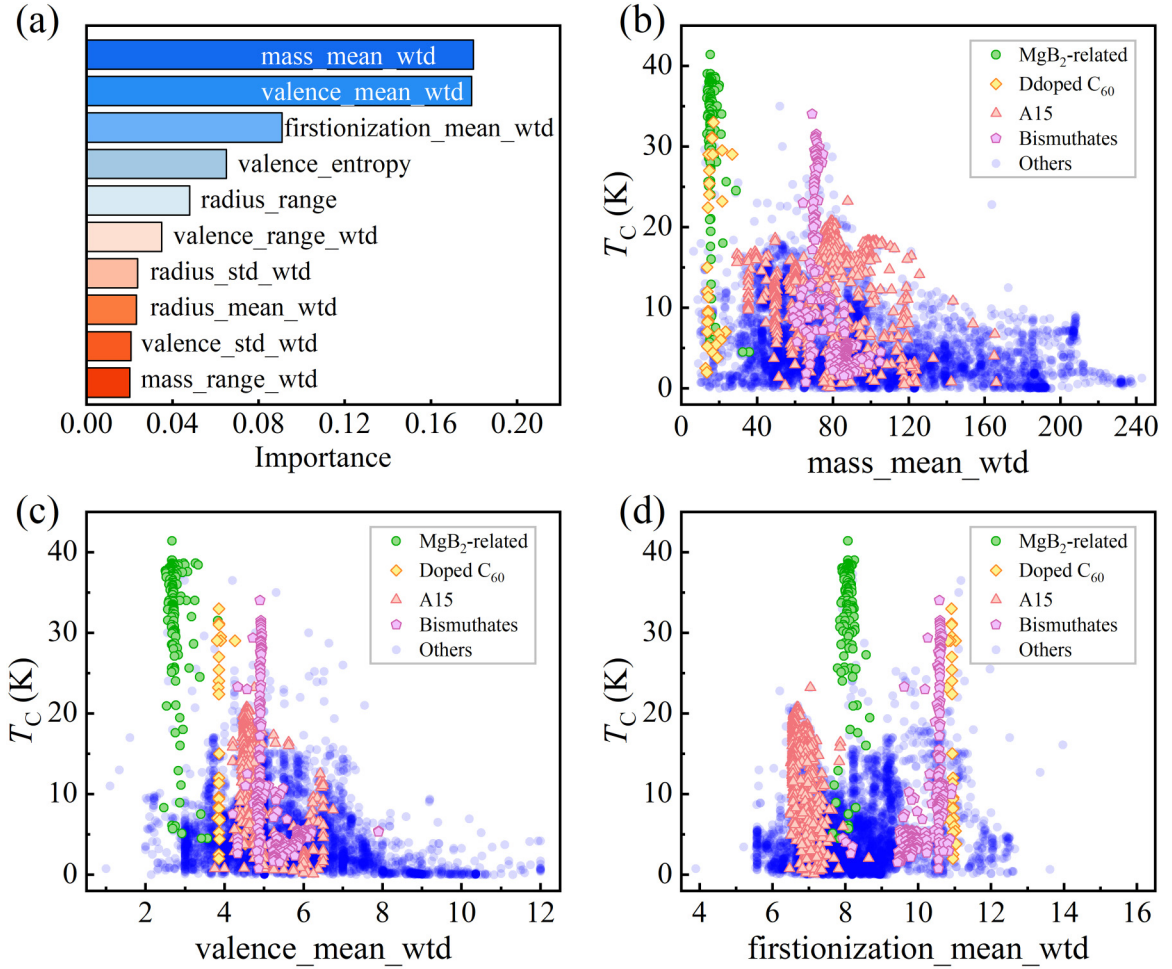


FIG. 3. (a) Top 10 influential features for conventional superconductors identified by the ML model. T_C is plotted vs the three most critical features: (b) weighted mean mass (mass_mean_wtd), (c) weighted mean valence (valence_mean_wtd), and (d) weighted mean first ionization energy (firstionization_mean_wtd) for conventional superconductors in specific families.

The weighted standard deviation of the radius (radius_std_wtd) feature quantifies the statistical distribution of atomic radii within a compound, reflecting the degree of atomic-scale disorder in the crystal lattice. As shown in Fig. 4(c), the impact of radius_std_wtd on T_C exhibits a trend strikingly similar to that of valence_std_wtd, peaking at approximately 0.6. This suggests that moderate variations in atomic size facilitate the emergence of high T_C . For the unfilled d electrons (d_{unfill}) feature in Fig. 4(d), T_C basically decreases as the value increases, consistent with a prior study [40]. It exhibits the highest T_C when d_{unfill} is in the range 0–1, correlating with the proportion of Cu-O layers: Compounds with high Cu-O contents are near this range, while doped systems with elements like La and Y (which have fewer d electrons) increase the average d_{unfill} value due to their abundant unfilled d orbitals.

C. ML predictions and DFT calculations

Given that materials design aims to unearth novel superconducting materials from the vast chemical landscape, we focused on the publicly accessible Materials Project database [41]. The database contains approximately 155 361 com-

pounds, each characterized by a wealth of computational properties, such as electronic energy bands, formation energies, and magnetic properties. This comprehensive dataset lays a solid foundation for the efficient screening of potential superconductors. Our methodology employed a rigorous multistage screening protocol, as illustrated in Fig. 1. The screening criteria were as follows.

In the first step, we set the following screening criteria: band gap = 0 eV, magnetic moment $m < 0.01 \mu_B/\text{f.u.}$, and formation energy $E_{\text{form}} < 0$ eV/atom. This step ensures that the candidate materials have metallic, nonmagnetic properties and thermodynamic stability. Based on the above criteria, we identified 23 635 candidate materials that meet the basic requirements for conventional superconductors. In the next step, we selected the binary and ternary compounds containing light B, C, and N elements, reducing the candidate materials to 2249. Furthermore, binary and ternary compounds with simpler structures (containing fewer than nine atoms in the unit cell) were further evaluated for their superconductivity by calculating the EPC at the DFT level.

Finally, we employed the ML-trained model to predict T_C of the 2249 previously identified candidates. We prioritized materials with higher T_C^{ML} ($T_C^{\text{ML}} \geq 10$ K) and smaller

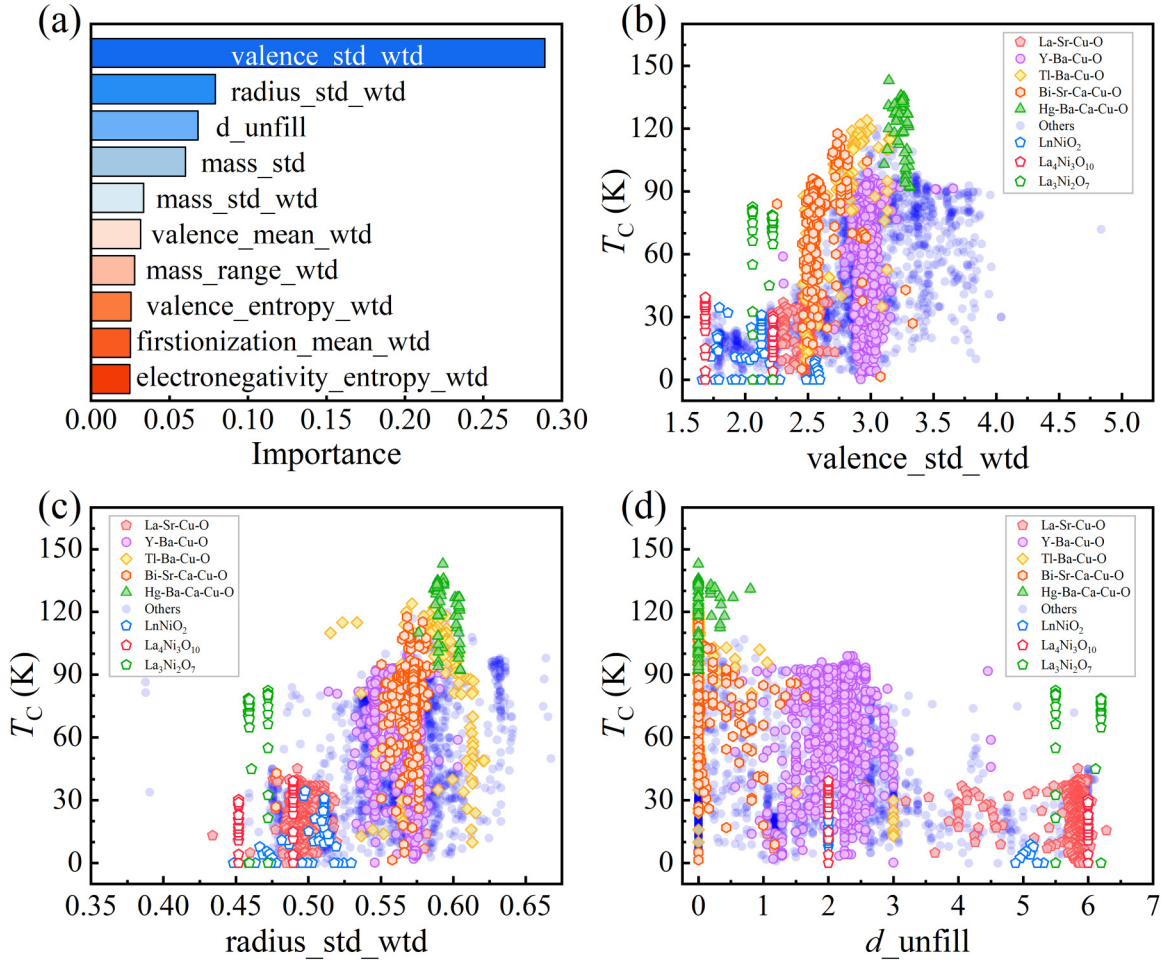


FIG. 4. (a) Top 10 influential features for high- T_C superconductors identified by the ML model. T_C is plotted vs the three most important features: (b) the weighted standard deviation of valence electrons (valence_std_wtd), (c) the weighted standard deviation of the atomic radius (radius_std_wtd), and (d) the number of unfilled d -orbital electrons (d_{unfill}) for cuprate and nickelate superconductors in specific families.

unit cells ($N_{\text{sites}} \leq 8$) for more accurate DFT calculations, selecting 100 materials (see Table SI of the Supplemental Material [18]). As listed in Table III, we ultimately identified 13 compounds with $T_C^{\text{DFT}} \geq 10$ K, including 3 compounds with $T_C^{\text{DFT}} \geq 25$ K, two structural forms of B₂CN ($T_C = 47.6$ and 43.7 K) and TiNbN₂ ($T_C = 25.3$ K). It is noteworthy that B₂CN [42], CaB₂ [8,43,44], and KB₆ [9] were also theoretically predicted as potential superconductors in previous DFT calculations. Among the predicted B₂CN polymorphs, the structure with space group $P\bar{4}m2$ (mp-1008527) has a large absolute value of formation energy ($|E_{\text{form}}| = 0.44$ eV/atom), indicating it is the most stable in thermodynamics and thus the most likely to be synthetically accessible.

For conventional superconductors, EPC serves as the key mechanism for the formation of superconducting Cooper pairs. Here, we systematically calculate the phonon dispersion and EPC of the 100 ML-predicted B-C-N based superconducting candidates via DFT calculations. We show the results for four typical materials in Fig. 5 (see more results in Figs. S7–S9 of the Supplemental Material [18]). These compound structures are dynamically stable due to the absence of imaginary (negative) phonon frequencies. When examining the Eliashberg function, we observe that the electron-phonon in-

teractions are dominant for soft phonon modes. For example, the EPC of MgB₂ is predominantly distributed at higher frequencies, whereas the EPC of B₂CN exhibits significant strength across both low and intermediate frequencies, resulting in its notably high T_C .

We further investigated the electronic structures of the ML-predicted B-C-N based superconducting candidates. We found that most compounds have a similar feature with strong σ bonds. For example, as illustrated in Fig. 6, the orbital-projected band structures of MgB₂ and CaB₂ show that B atoms adopt sp^2 hybridization, forming in-plane σ bonds (head-on p_x and p_y orbital overlap) and vertical π bonds (side-by-side p_z orbital overlap). The σ -bonding bands directly cross the Fermi energy level and become the main source of carriers for superconducting pairing. However, KB₆ and B₂CN feature σ -bonding bands formed via sp^3 hybridization (B-B and B-C-N bonds), resulting in strong σ -bonding bands. The mechanism of metallization of strong σ bonds can contribute to high T_C values, as pointed out in previous studies [45–47].

Next, we investigated the potential high- T_C candidates in La₃Ni₂O₇-type bilayer nickelates. Using the ML-trained model, we predicted T_C of $R_3\text{Ni}_2\text{O}_7$ (R = rare-earth elements from La to U) nickelates under a pressure of 30 GPa, at

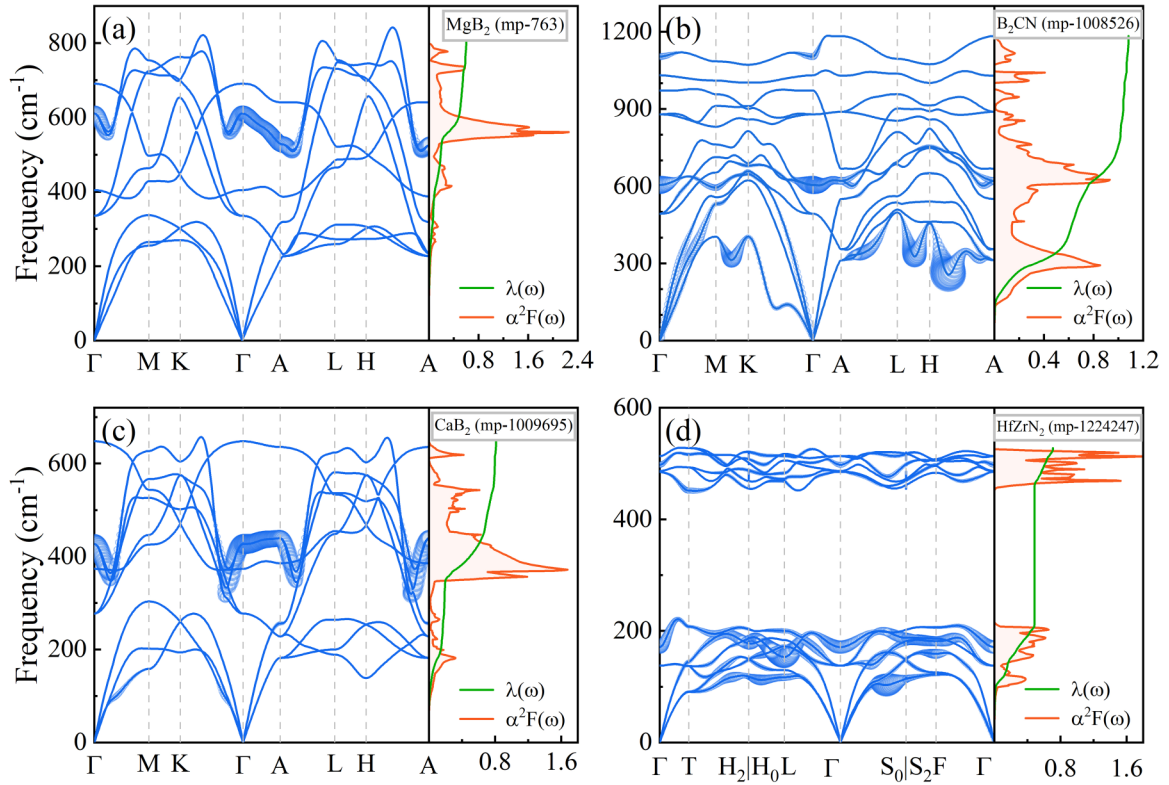


FIG. 5. Left: Phonon spectra of (a) MgB_2 (mp-763), (b) B_2CN (mp-1008525), (c) CaB_2 (mp-1009695), and (d) HfZrN_2 (mp-1224247). The size of the blue circle is proportional to the magnitude of the electron-phonon coupling strength (λ_{qv}). Right: The Eliashberg spectral function $\alpha^2F(\omega)$ (orange line) and $\lambda(\omega)$ (green line).

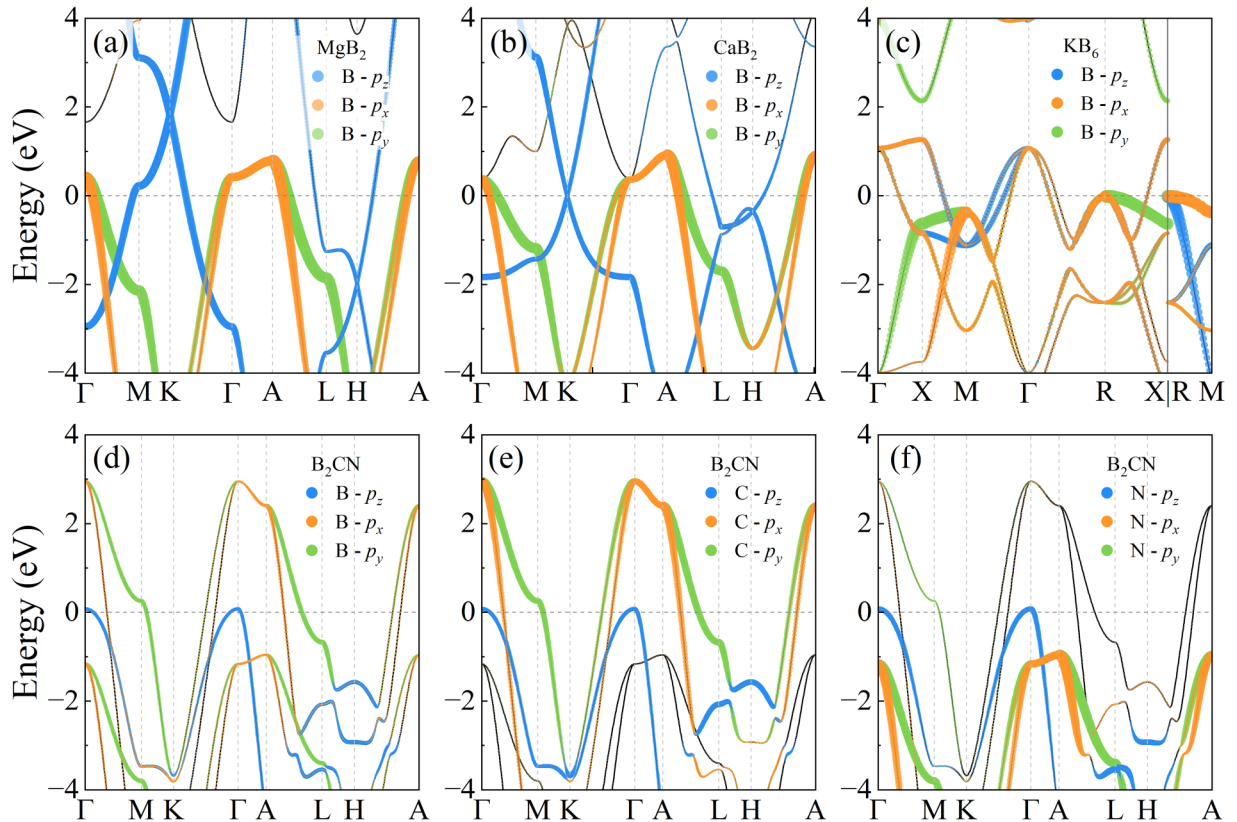


FIG. 6. Orbital-projected electronic band structures of (a) MgB_2 , (b) CaB_2 , (c) KB_6 , and (d)–(f) B_2CN . The widths of the blue, orange, and green lines are proportional to the weights of the p_z , p_x , and p_y orbitals, respectively.

TABLE III. Predicted superconductor candidates with $T_C^{\text{ML}} \geq 10$ K and $T_C^{\text{DFT}} \geq 5$ K. The Materials Project ID number (mp-id), formula, space group, ML-predicted T_C^{ML} , DFT-predicted T_C^{DFT} , and ID number in the Inorganic Crystal Structure Database (ICSD) are shown. T_C^{DFT} was obtained by using the Allen-Dynes-modified McMillan equation with the typical Coulomb pseudopotential $\mu^* = 0.10$.

mp-id	Formula	Space group	T_C^{ML} (K)	T_C^{DFT} (K)	ICSD
mp-1008525	B ₂ CN	<i>R3m</i>	16.3	47.6	183792
mp-1008526	B ₂ CN	<i>P3m1</i>	16.3	43.7	183791
mp-35869	TiNbN ₂	<i>I4₁/amd</i>	15.7	25.3	
mp-1009695	CaB ₂	<i>P6/mmm</i>	9.7	21.4	186764
mp-1008527	B ₂ CN	<i>P4₂m2</i>	16.3	19.9	183790
mp-1216616	V ₂ CN	<i>R3$\bar{m}$</i>	14.8	16.7	
mp-1215188	ZrTiN ₂	<i>R3$\bar{m}$</i>	12.2	15.7	
mp-1224257	HfTiN ₂	<i>P4/mmm</i>	11.7	14.5	
mp-1224286	HfTiN ₂	<i>R3$\bar{m}$</i>	11.7	14.2	
mp-1215180	ZrTiN ₂	<i>P4/mmm</i>	12.2	13.3	
mp-37179	Ta ₂ CN	<i>I4₁/amd</i>	9.6	12.7	
mp-1224247	HfZrN ₂	<i>R3$\bar{m}$</i>	12.7	12.1	
mp-1220685	Nb ₂ Mo ₂ C ₃	<i>R3$\bar{m}$</i>	9.7	10.6	
mp-1217026	TiB ₄ Mo	<i>P6/mmm</i>	10.9	9.4	43665
mp-1215387	Zr ₄ CN ₃	<i>R3$\bar{m}$</i>	14.0	9.2	
mp-1217107	Ti ₄ CN ₃	<i>R3$\bar{m}$</i>	12.9	9.2	
mp-1224384	Hf ₄ CN ₃	<i>R3$\bar{m}$</i>	10.0	6.5	
mp-1076	KB ₆	<i>Pm3$\bar{m}$</i>	25.1	8.8	98990
mp-2646985	V ₂ GaN	<i>P6₃/mmc</i>	9.5	7.7	
mp-1246929	Zr ₄ BN ₃	<i>Pm3$\bar{m}$</i>	12.5	5.4	

which the experimentally observed T_C of La₃Ni₂O₇ reaches a maximum value of 80 K. As shown in Fig. 7(a), T_C exhibits nonmonotonic variation with an increasing atomic number of

R. La₃Ni₂O₇ demonstrates the highest T_C , while Tb₃Ni₂O₇ and Ac₃Ni₂O₇ show locally maximal T_C values. Interestingly, our ML-identified Tb₃Ni₂O₇ and Ac₃Ni₂O₇ were also

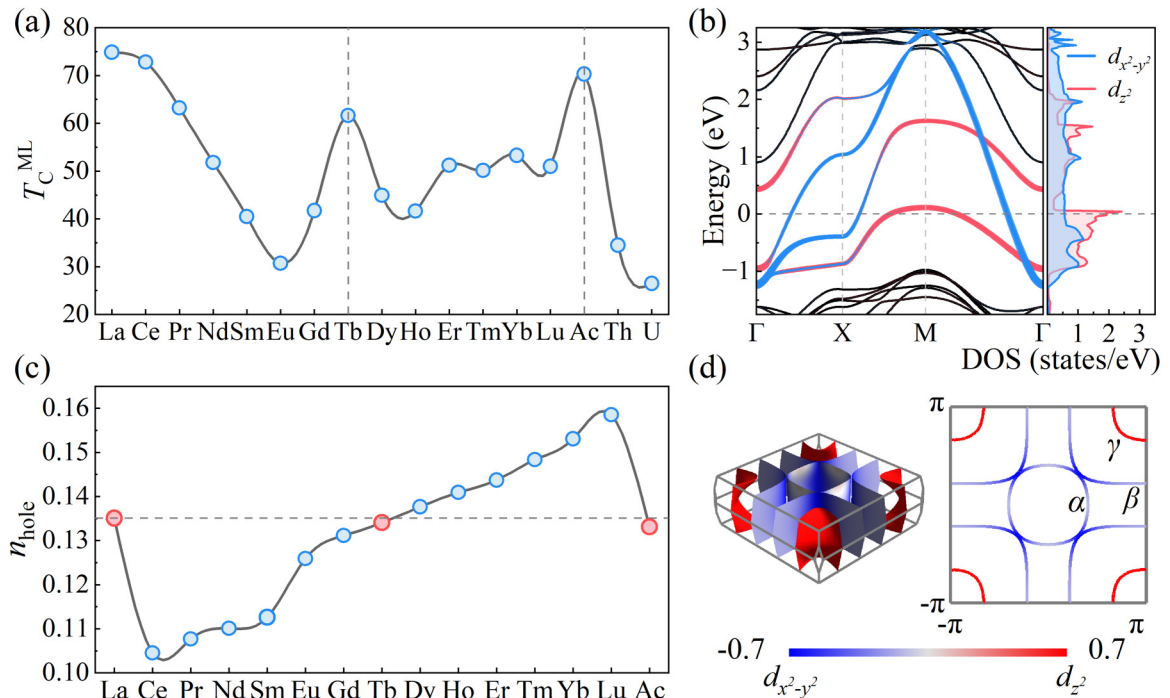


FIG. 7. (a) Predicted T_C of the $R_3\text{Ni}_2\text{O}_7$ nickelates under 30 GPa pressure using the ML-trained model. T_C values of Tb₃Ni₂O₇ and Ac₃Ni₂O₇ are local maxima. (b) Orbital-projected band structure and density of states of La₃Ni₂O₇ at 30 GPa. The widths of the red and blue lines are proportional to the weight of Ni $3d_{z^2}$ and Ni $3d_{x^2-y^2}$ orbitals, respectively. (c) Hole concentrations of the γ band dominated by Ni $3d_{z^2}$ in $R_3\text{Ni}_2\text{O}_7$. (d) Fermi surfaces in the three-dimensional Brillouin zone and projection on the two-dimensional k_x - k_z plane with orbital weights of La₃Ni₂O₇ under 30 GPa. The three pockets are labeled as α , β , and γ , respectively.

proposed as potential high- T_C superconducting nickelates in previous theoretical studies [48–50].

Given that nickelates are strongly correlated systems and belong to unconventional superconductors whose behavior deviates from the BCS theory, one cannot directly calculate their T_C using the DFT method. Instead, we adopted a comparative approach by analyzing the key features of the electronic structure of $R_3\text{Ni}_2\text{O}_7$. The DFT-calculated bands and Fermi surfaces of $R_3\text{Ni}_2\text{O}_7$ are similar to those of $\text{La}_3\text{Ni}_2\text{O}_7$ (see more results in Figs. S10–S15 of the Supplemental Material [18]). Moreover, we also calculated band structures of $R_3\text{Ni}_2\text{O}_7$ by considering the $R4f$ electrons as valence electrons to examine the influence of the $4f$ states on the electronic structure. For most R elements (excluding Ce, Sm, and Tm), the $R4f$ states are energetically separated from the Fermi energy level and exhibit negligible hybridization with the relevant Ni $3d$ bands (see Figs. S16–S19 and further discussion in Sec. 7 of the Supplemental Material [18]). As shown in Figs. 7(b) and 7(d), three bands cross the Fermi energy level: two electron-type α and β bands primarily contributed by the Ni $3d_{x^2-y^2}$ orbitals and one hole-type γ band mainly contributed by the Ni $3d_z^2$ orbital. Although the band structures of $R_3\text{Ni}_2\text{O}_7$ are similar, a more detailed analysis revealed that the hole-type γ band varies depending on the R element.

To quantify the variation, we calculated the hole-carrier concentrations contributed by the γ pocket, shown in Fig. 7(c). The hole concentration exhibits a rapid decrease from La to Ce, followed by a gradual increase with rising R atomic number and then a sharp drop upon reaching Ac. This trend remarkably coincides with the variation of valence_std_wtd across different $R_3\text{Ni}_2\text{O}_7$ compounds (see Fig. S6 of the Supplemental Material [18]). Note that for Th and U, the treatment of $4f$ electrons in pseudopotentials as valence electrons leads to poor SCF convergence in DFT calculations, yielding unreliable results that are therefore not presented. Surprisingly, the hole concentrations n_{hole} of γ pockets for $\text{Tb}_3\text{Ni}_2\text{O}_7$ and $\text{Ac}_3\text{Ni}_2\text{O}_7$ are 0.134 and 0.133, which are very close to the value of 0.135 for $\text{La}_3\text{Ni}_2\text{O}_7$. Previous theoretical calculations proposed that the γ pocket contributed by the Ni $3d_z^2$ orbital plays a dominant role in the superconductivity of $\text{La}_3\text{Ni}_2\text{O}_7$, determining the pairing symmetry as $s^\pm$ [4,24,51–

53]. The γ bands of $\text{Tb}_3\text{Ni}_2\text{O}_7$ and $\text{Ac}_3\text{Ni}_2\text{O}_7$ identified exhibit strong consistency with those of $\text{La}_3\text{Ni}_2\text{O}_7$, suggesting their potential as bilayer nickelate superconductors.

IV. CONCLUSIONS

In this study, we developed a large-scale strategy to accelerate the discovery of superconductors by combining data-driven ML techniques and first-principles calculations. We systematically explored binary and ternary B-C-N based conventional superconductors as well as potential high- T_C $R_3\text{Ni}_2\text{O}_7$ nickelates. From an initial pool of over 150 000 materials, our ML model identified 100 binary and ternary B-C-N based compounds with $T_C^{\text{ML}} \geq 10$ K. Among these, 13 compounds were further validated using DFT, yielding $T_C^{\text{DFT}} \geq 10$ K, including 3 compounds with $T_C^{\text{DFT}} \geq 25$ K. The predicted B-C-N based compounds exhibit strong σ -bonding interactions, similar to those of MgB_2 . Furthermore, our approach uncovered two promising nickelate superconducting candidates, $\text{Tb}_3\text{Ni}_2\text{O}_7$ and $\text{Ac}_3\text{Ni}_2\text{O}_7$, which demonstrate high T_C values comparable to $\text{La}_3\text{Ni}_2\text{O}_7$. Computational analyses revealed that these nickelates share a hole-type band derived from Ni $3d_z^2$ orbitals, accompanied by comparable hole carrier concentrations. Critical feature analysis underscores key descriptors responsible for both conventional and high- T_C superconductors, facilitating not only the efficient discovery of novel materials but also deeper mechanistic insights. This work demonstrates a robust pathway for predicting both conventional and high- T_C superconducting materials and contributes to a more comprehensive understanding of their underlying mechanisms.

ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China (Grant No. 2022YFA1403200). The calculations were performed at Hefei Advanced Computing Center.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

-
- [1] H. K. Onnes, The resistance of pure mercury at helium temperatures, *Commun. Phys. Lab. Univ. Leiden* **12**, 120 (1911).
 - [2] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Theory of superconductivity, *Phys. Rev.* **108**, 1175 (1957).
 - [3] J. Nagamatsu, N. Nakagawa, T. Muranaka, Y. Zenitani, and J. Akimitsu, Superconductivity at 39 K in magnesium diboride, *Nature (London)* **410**, 63 (2001).
 - [4] H. L. Sun, M. W. Huo, X. W. Hu, J. Y. Li, Z. J. Liu, Y. F. Han, L. Y. Tang, Z. Q. Mao, P. T. Yang, B. S. Wang, J. G. Cheng, D. X. Yao, G. M. Zhang, and M. Wang, Signatures of superconductivity near 80 K in a nickelate under high pressure, *Nature (London)* **621**, 493 (2023).
 - [5] C. Park and R. L. Snyder, Structures of high-temperature cuprate superconductors, *J. Am. Ceram. Soc.* **78**, 3171 (1995).
 - [6] R. Kitchin, Big data, new epistemologies and paradigm shifts, *Big Data Soc.* **1**, 1 (2014).
 - [7] G. R. Schleder, A. C. M. Padilha, C. M. Acosta, M. Costa, and A. Fazzio, From DFT to machine learning: Recent approaches to materials science — a review, *J. Phys. Mater.* **2**, 032001 (2019).
 - [8] Z. Yu, T. Bo, B. Liu, Z. Fu, H. Wang, S. Xu, T. Xia, S. Li, S. Meng, and M. Liu, Superconductive materials with MgB_2 -like structures from data-driven screening, *Phys. Rev. B* **105**, 214517 (2022).

- [9] K. Hamidieh, A data-driven statistical model for predicting the critical temperature of a superconductor, *Comput. Mater. Sci.* **154**, 346 (2018).
- [10] H. Gashmard, H. Shakeripour, and M. Alaei, Predicting superconducting transition temperature through advanced machine learning and innovative feature engineering, *Sci. Rep.* **14**, 3965 (2024).
- [11] M. R. Quinn and T. M. McQueen, Identifying new classes of high temperature superconductors with convolutional neural networks, *Front. Electron. Mater.* **2**, 893797 (2022).
- [12] N. Hoffmann, T. F. T. Cerqueira, J. Schmidt, and M. A. L. Marques, Superconductivity in antiperovskites, *npj Comput. Mater.* **8**, 150 (2022).
- [13] N. Hoffmann, T. F. T. Cerqueira, P. Borlido, A. Sanna, J. Schmidt, and M. A. L. Marques, Searching for ductile superconducting Heusler X_2YZ compounds, *npj Comput. Mater.* **9**, 138 (2023).
- [14] L. Gu, Y. Liu, P. Chen, H. Y. Huang, N. Chen, Y. Li, T. Lookman, Y. T. Lu, and Y. J. Su, Bond sensitive graph neural networks for predicting high temperature superconductors, *MGE Adv.* **2**, e48 (2024).
- [15] B. W. Jiang, X. S. Luo, Y. Sun, X. Zhong, J. Lv, Y. Xie, Y. M. Ma, and H. Y. Liu, Data-driven search for high-temperature superconductors in ternary hydrides under pressure, *Phys. Rev. B* **111**, 054505 (2025).
- [16] L. L. Zha, J. J. Jiang, Y. M. Xue, Z. B. Cheng, S. W. Yao, W. J. Hu, L. Peng, T. T. Shi, J. Chen, X. L. Liu, and J. Lin, Stability and superconductivity of hexagonal prism-structured polyhydrides X_2MgH_{18} ($X = Li, Na, K, Rb, Cs$) under moderate pressure, *Phys. Chem. Chem. Phys.* **27**, 3844 (2025).
- [17] Y. Du, Z. F. Wang, H. Y. Liu, G. J. Liu, and X. Zhong, Superconductivity of electron-doped chalcogenides under high pressure, *Phys. Rev. Res.* **7**, 013049 (2025).
- [18] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/wn98-9yc7> for the details of the machine learning accelerated search for superconductors in B-C-N based compounds and $R_3Ni_2O_7$ -type nickelates.
- [19] Materials Database Group, MDR SuperCon Datasheet, Ver.220808 (National Institute for Materials Science, 2011), <https://doi.org/10.48505/nims.3837>.
- [20] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [21] P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* **50**, 17953 (1994).
- [22] P. Giannozzi, *et al.*, QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [23] J. Noffsinger, F. Giustino, B. D. Malonea, C. H. Park, S. G. Louie, and M. L. Cohen, EPW: A program for calculating the electron-phonon coupling using maximally localized Wannier functions, *Comput. Phys. Commun.* **181**, 2140 (2010).
- [24] Z. H. Luo, X. W. Hu, M. Wang, W. Wu, and D. X. Yao, Bilayer two-orbital model of $La_3Ni_2O_7$ under pressure, *Phys. Rev. Lett.* **131**, 126001 (2023).
- [25] P. B. Allen and R. C. Dynes, Transition temperature of strong-coupled superconductors reanalyzed, *Phys. Rev. B* **12**, 905 (1975).
- [26] J. Li, X. P. Cui, Y. Jin, X. Q. Yin, S. X. Cao, Z. J. Feng, and J. C. Zhang, Appearance of superconductivity at 15.9 K in layered V_2AlN , *J. Supercond. Novel Magn.* **30**, 1 (2017).
- [27] J. N. Graham, H. Liu, V. Sazgari, C. Mielke, III, M. Medarde, H. Luetkens, R. Khasanov, Y. Shi, and Z. Guguchia, Microscopic probing of the superconducting and normal state properties of $Ta_2V_{3.1}Si_{0.9}$ by muon spin rotation, *Commun. Mater.* **5**, 225 (2024).
- [28] Y. J. Lee, O. Inturu, J. H. Kim, and J. S. Rhyee, Robust bulk superconductivity by giant proximity effect in Weyl semimetal-superconducting NbP/NbSe₂ composites, *J. Appl. Phys.* **132**, 123903 (2022).
- [29] H. Ishikawa, U. Wedig, J. Nuss, R. K. Kremer, R. Dinnebier, M. Blankenhorn, M. Pakdaman, Y. Matsumoto, T. Takayama, K. Kitagawa, and H. Takagi, Superconductivity at 4.8 K and violation of Pauli limit in La_2IRu_2 comprising Ru honeycomb layer, *Inorg. Chem.* **58**, 12888 (2019).
- [30] D. Tay, T. Shang, Y. P. Qi, T. P. Ying, H. Hosono, H. R. Ott, and T. Shiroka, *s*-wave superconductivity in the noncentrosymmetric W_3Al_2C superconductor: An NMR study, *J. Phys.: Condens. Matter* **34**, 194005 (2022).
- [31] K. W. Chen, G. Chappell, S. Zhang, W. Lan, T. Besara, K. Huang, D. Graf, L. Balicas, A. P. Reyes, and R. E. Baumbach, Superconductivity in single crystals of $ZrP_{1.27}Se_{0.73}$, *Phys. Rev. B* **102**, 144522 (2020).
- [32] B. Zheng, X. K. Feng, B. Liu, Z. F. Liu, S. S. Wang, Y. Zhang, X. Ma, Y. Luo, C. L. Wang, R. M. Li, Z. Y. Zhang, S. T. Cui, Y. L. Lu, Z. Sun, J. F. He, S. Y. A. Yang, and B. Xiang, The coexistence of superconductivity and topological order in van der Waals $InNbS_2$, *Small* **20**, 2305909 (2024).
- [33] Z. H. Yang, Z. Yang, Q. P. Su, E. D. Fang, J. H. Yang, B. Chen, H. D. Wang, J. H. Du, C. X. Wu, and M. H. Fang, Superconductivity in $TlBi_2$ with a large Kadowaki-Woods ratio, *Phys. Rev. B* **106**, 224501 (2022).
- [34] T. Shang, D. Tay, H. Su, H. Q. Yuan, and T. Shiroka, Evidence of fully gapped superconductivity in $NbReSi$: A combined μ SR and NMR study, *Phys. Rev. B* **105**, 144506 (2022).
- [35] Q. Q. Zhu, G. R. Xiao, W. Z. Yang, S. J. Song, G. H. Cao, and Z. Ren, Mo_3ReRuC : A Noncentrosymmetric Superconductor Formed in the $MoReRu-Mo_2C$ System, *Inorg. Chem.* **61**, 17115 (2022).
- [36] Y. Liu, J. Li, W. Z. Yang, J. Y. Lu, B. Y. Cao, H. X. Li, W. L. Chai, S. Q. Wu, B. Z. Li, Y. L. Sun, W. H. Jiao, C. Wang, X. F. Xu, Z. Ren, and G. H. Cao, Superconductivity in Kagome metal $ThRu_3Si_2$, *Chin. Phys. B* **33**, 057401 (2024).
- [37] C. A. Reynolds, B. Serin, W. H. Wright, and L. B. Nesbitt, Superconductivity of isotopes of mercury, *Phys. Rev.* **78**, 487 (1950).
- [38] M. Pop, G. Borodi, and S. Simon, Correlation between valence electron concentration and high-temperature superconductivity, *J. Phys. Chem. Solids* **61**, 1939 (2000).
- [39] B. T. Matthias, Empirical relation between superconductivity and the number of valence electrons per atom, *Phys. Rev.* **97**, 74 (1955).
- [40] V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, and I. Takeuchi, Machine learning modeling of superconducting critical temperature, *npj Comput. Mater.* **4**, 29 (2018).
- [41] S. P. Ong, S. Cholia, A. Jain, M. Brafman, D. Gunter, G. Ceder, and K. A. Persson, The materials application programming interface (API): A simple, flexible, and efficient API for materials

- data based on representational state transfer (REST) principles, *Comput. Mater. Sci.* **97**, 209 (2015).
- [42] Q. Li, D. Zhou, H. Wang, W. J. Chen, B. J. Wu, Z. J. Wu, and W. T. Zheng, Crystal and electronic structures of superhard B₂CN: An *ab initio* study, *Solid State Commun.* **152**, 71 (2012).
- [43] N. I. A. Ahmed and C. Parlak, First-principles calculations on mechanical and electronic properties of AlB₂-type CaB₂, *Mater. Today Commun.* **38**, 108290 (2024).
- [44] H. J. Choi, S. G. Louie, and M. L. Cohen, Prediction of superconducting properties of CaB₂ using anisotropic Eliashberg theory, *Phys. Rev. B* **80**, 064503 (2009).
- [45] M. Gao, Z. Y. Lu, and T. Xiang, Finding high-temperature superconductors by metallizing the σ -bonding electrons, *Physics* **44**, 421 (2015).
- [46] J. M. An and W. E. Pickett, Superconductivity of MgB₂: Covalent bonds driven metallic, *Phys. Rev. Lett.* **86**, 4366 (2001).
- [47] Y. Kong, O. V. Dolgov, O. Jepsen, and O. K. Andersen, Electron-phonon interaction in the normal and superconducting states of MgB₂, *Phys. Rev. B* **64**, 020501(R) (2001).
- [48] B. Geisler, J. J. Hamlin, G. R. Stewart, R. G. Hennig, and P. J. Hirschfeld, Structural transitions, octahedral rotations, and electronic properties of A₃Ni₂O₇ rare-earth nickelates under high pressure, *npj Quantum Mater.* **9**, 38 (2024).
- [49] S. Q. Wu, Z. H. Yang, X. Ma, J. H. Dai, M. Shi, H. Q. Yuan, H. Q. Lin, and C. Cao, Ac₃Ni₂O₇ and La₂AeNi₂O₆F (Ae = Sr, Ba): Benchmark materials for bilayer nickelate superconductivity, [arXiv:2403.11713](https://arxiv.org/abs/2403.11713).
- [50] L. C. Rhodes and P. Wahl, Structural routes to stabilize superconducting La₃Ni₂O₇ at ambient pressure, *Phys. Rev. Mater.* **8**, 044801 (2024).
- [51] Q. G. Yang, D. Wang, and Q. H. Wang, Possible $s^{\pm}$ wave superconductivity in La₃Ni₂O₇, *Phys. Rev. B* **108**, L140505 (2023).
- [52] Y. H. Gu, C. C. Le, Z. S. Yang, X. X. Wu, and J. P. Hu, Effective model and pairing tendency in the bilayer Ni-based superconductor La₃Ni₂O₇, *Phys. Rev. B* **111**, 174506 (2025).
- [53] Z. Y. Shao, Y. B. Liu, M. Liu, and F. Yang, Band structure and pairing nature of La₃Ni₂O₇ thin film at ambient pressure, *Phys. Rev. B* **112**, 024506 (2025).