J Adv Ceram

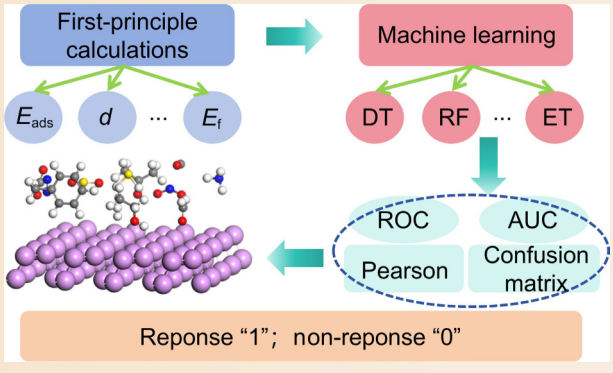
Research Article

# First-principles calculations informing machine learning framework and visualization system for rapid and generalized gas response prediction in black phosphorus sensors

Mingyuan Wang[1], Yaqi Zhang[2], Bowen Xiong[1], Ke Wang[3], Xiangzhao Zhang[2], Jian Yang[2], Lin Xu[1], Guanjun Qiao[2], Guiwu Liu[2]✉

Cite this article: Wang M, Zhang Y, Xiong B, et al. *J Adv Ceram* 2026, **15**(3): 9221243. https://doi.org/10.26599/JAC.2026.9221243

**ABSTRACT:** Gas sensors are vital in practical applications, yet the efficient screening of sensing materials remains a formidable challenge. Conventional trial-and-error approaches are costly, single descriptors fail to capture complex interactions, and multiparameter combinations introduce nonlinearities. To overcome these limitations, we propose a synergistic strategy that integrates first-principles calculations with machine learning (ML) for the rapid prediction of gas sensitivity. Using black phosphorus (BP) as a model system, we evaluated its responsiveness to 21 gases by analyzing adsorption-induced electronic and structural changes. Key descriptors extracted from these calculations were used to train six ML models. The extra tree (ET) model demonstrated exceptional robustness, achieving 96% accuracy with minimal deviation in fivefold cross-validation and top-tier performance in F1-score evaluation. Furthermore, analyses of feature importance and SHapley Additive exPlanations (SHAP) identified the adsorption energy, p-orbital center, valence band maximum, conduction band minimum, and Fermi level as the dominant descriptors. We also developed a lightweight, Python-based prediction and visualization system. By inputting only these five key features obtained from first-principles calculations, this tool enables real-time assessment of BP's response to various gas molecules. This integrated approach demonstrated significant potential for predicting material sensing properties and offers valuable theoretical guidance for designing gas sensors.




**KEYWORDS:** black phosphorus (BP); gas response; first-principles calculations; machine learning (ML); visualization system

## 1 Introduction

With the rapid advancement of industrialization, large quantities of toxic and hazardous gases such as $NO_2$, CO, and $NH_3$ are emitted during production processes, posing serious threats to human health and the ecological environment [1,2]. Consequently, the real-time and efficient detection of gas species and concentrations in the environment is crucial for ensuring the safety of both industrial production and daily life [3,4]. Among various gas detection technologies, gas sensors have emerged as one of the most widely used approaches due to their simple structure, fast response, and high sensitivity. Gas sensors have found broad applications in industrial safety, environmental monitoring, medical diagnostics, and smart home systems [5,6]. Currently, the mainstream gas-sensitive materials are metal oxide semiconductors (MOSs), such as $SnO_2$, $In_2O_3$, ZnO, and $WO_3$ [7–9]. These materials are extensively studied and applied due to their low cost, rapid response and recovery time, and ease of fabrication [10]. However, MOS-based gas sensors typically require higher operating temperatures (> 200 °C) to achieve satisfactory sensing performance, which not only increases energy consumption but also limits their use in wearable devices, portable sensors, and low-temperature environments [1]. To address the limitations of high-temperature operation, emerging two-dimensional (2D) materials, such as graphene, black phosphorus (BP), transition metal dichalcogenides (TMDs), and MXenes, have gained significant attention for room-temperature gas sensing owing to their large surface area, high carrier mobility,

[1] School of Mechanical Engineering, Jiangsu University, Zhenjiang 212013, China. [2] School of Materials Science and Engineering, Jiangsu University, Zhenjiang 212013, China. [3] Beijing Key Laboratory of Ionic Liquids Clean Process, CAS Key Laboratory of Green Process and Engineering, State Key Laboratory of Multiphase Complex Systems, Innovation Academy for Green Manufacture, Institute of Process Engineering, Chinese Academy of Sciences, Beijing 100190, China.

✉ Corresponding author. E-mail: gwliu76@ujs.edu.cn

Received: November 11, 2025; Revised: December 9, 2025; Accepted: January 5, 2026

and excellent ambient-condition responsiveness [11,12]. For instance, a larger specific surface area provides a greater contact interface between the material and gas molecules, which facilitates stronger adsorption and enhances sensitivity; higher carrier mobility allows the material's electrical response to occur more rapidly, thereby improving the response speed of the gas sensor. Among these 2D materials, BP, a layered 2D material with semiconducting properties and outstanding electronic transport characteristics, has shown great potential for gas detection applications [13,14]. Notably, BP prepared via liquid phase exfoliation under a nitrogen atmosphere exhibits superior $NO_2$ sensing performance at room temperature compared to other 2D materials. Studies have shown that BP can detect $NO_2$ at concentrations as low as 0.1 ppm, a threshold beyond the capability of $MoS_2$ and reduced graphene oxide (rGO) [15]. Moreover, BP's response speed to $NO_2$ is approximately 40 times faster than these materials, underscoring its outstanding advantages for high-sensitivity detection of low-concentration gases [15]. Therefore, due to its excellent sensing performance and room temperature operability, BP has become a highly promising candidate for the next generation of gas sensing materials, particularly for the rapid and precise detection of low concentrations of toxic gases.

Currently, the selection of gas sensing materials still largely relies on traditional "trial-and-error" approaches, which typically require extensive experimental validation. This process is typically resource-intensive, requiring significant time, effort, and expense [16]. Second, although first-principles calculations can predict key parameters at the atomic scale (such as adsorption energy, adsorption distance, and charge transfer) and thereby provide important theoretical insights into gas–material interactions, this process typically relies on a single descriptor, which often fails to capture the full complexity of these interactions, potentially leading to biased or incomplete predictions. Combining multiple parameters offers a more comprehensive representation of the structural and electronic characteristics of materials during gas adsorption and response, but it also introduces highly complex nonlinear relationships, making it difficult to establish unified and generalizable screening criteria. This greatly undermines the effectiveness and scalability of traditional approaches in practical applications. More critically, most studies on gas response prediction are limited to small-scale datasets with restricted volume and representativeness, lacking intelligent methods for large-scale system analysis and rapid screening. Such limitations restrict the generalizability of models across diverse material systems and diminish their practical utility in the high-throughput discovery and optimization of gas sensing materials. Thus, there is an urgent need to establish an efficient and accurate screening strategy to rapidly assess the responsiveness of materials to target gases. In this context, high-throughput screening has emerged as a promising solution, offering the ability to evaluate large-scale material systems within a short period and significantly accelerating the discovery of new sensing materials [17]. First-principles calculations have become a reliable theoretical tool in this process. Without the need for experimental input, it can efficiently predict key physicochemical properties of materials at the atomic scale, such as adsorption energy, adsorption distance, and charge transfer [18]. These parameters are commonly regarded as important indicators of the interaction strength between gas molecules and material surfaces and are thus widely used for evaluating gas sensing performance [19,20]. As a result, integrating multidimensional descriptors derived from first-principles calculations with advanced data-driven approaches has become a key research direction for the intelligent design and rapid screening of high-performance gas sensing materials.

To facilitate rapid and high-throughput prediction of material gas responses, we establish a theoretical framework and a lightweight Python-based visualization system integrating first-principles calculations with machine learning, capable of quickly predicting the responses of BP to diverse gases. The foundation of this research is based on gas sensing response data of BP toward 21 gases obtained from previous studies through both traditional experimental methods and theoretical simulations. These data are used to systematically evaluate the gas sensing behavior of BP at room temperature, which were sourced from multiple independent literature sources, where the testing procedures, environmental conditions, and data processing methods were inconsistent. Therefore, when constructing the experimental dataset, we only performed a binary classification of "response" versus "nonresponse" and did not further account for variations in sample quality, layer thickness, or defect density to avoid introducing uncontrollable biases. The types of gases and their corresponding response behaviors are detailed in Table 1, where "1" indicates that BP responds to the gas, and "0" indicates nonresponse. Here, simplifying the experimental observations into a binary response/nonresponse classification is a reasonable approach to maintain data consistency and avoid introducing uncontrollable systematic bias. With limited data, a binary classification model is more robust and better suited for capturing trend-level information. The primary purpose of using 0/1 labels in this study is to establish a fundamental framework for exploring the qualitative mapping between first-principles computational descriptors and experimental sensing responses. This simplification reduces the complexity of the model and helps verify the effectiveness of feature selection. However, this simplification has certain limitations: It cannot reflect response magnitude, detection limits, or kinetic behavior, nor can it capture quantitative variations arising from different experimental conditions. Moreover, first-principles calculations were employed to obtain a series of microstructural and electronic properties of BP after gas adsorption, including adsorption energy, adsorption distance, charge transfer, and electric properties. These parameters were used as input features for the machine learning model, while the experimentally observed gas response behavior served as the target values. Consequently, a classification model based on the ET algorithm was developed to predict the gas sensing behavior of BP. The model achieved a high prediction accuracy of 96%, demonstrating the feasibility and effectiveness of the proposed theoretical framework for the rapid evaluation of gas sensing

**Table 1** Target gases and BP responds to them

| Gas type | Response | Ref. | Gas type | Response | Ref. |
|---|---|---|---|---|---|
| $NO_2$ | 1 | [21] | NO | 1 | [22] |
| $NH_3$ | 1 | [23] | $N_2$ | 0 | [24] |
| $H_2S$ | 1 | [25] | $H_2O$ | 1 | [26] |
| $H_2$ | 0 | [27] | $CO_2$ | 0 | [28] |
| CO | 0 | [28] | $CH_3OH$ | 1 | [27] |
| $CH_3CHO$ | 0 | [27] | $CH_3COCH_3$ | 1 | [29] |
| $CH_3CHOHCH_3$ | 0 | [30] | $CH_3CH_2OH$ | 0 | [30] |
| $CH_3CH_2CHO$ | 1 | [31] | $CCl_3H$ | 0 | [30] |
| $CCl_2H_2$ | 0 | [30] | $C_7H_8$ | 0 | [27] |
| $C_6H_6$ | 0 | [32] | $C_3H_9O_3P$ | 1 | [33] |
| $C_3H_9N$ | 1 | [34] | | | |

Note: "1" indicates that BP can respond to the gas, and "0" indicates nonresponse.

performance. Feature importance ranking and SHAP interpretation consistently indicate that the adsorption energy ($E_{ads}$), p-orbital center ($\varepsilon_p$), valence band maximum (VBM), conduction band minimum (CBM), and Fermi level ($E_f$) play decisive roles in prediction, revealing the critical role of the synergistic regulation of the adsorption energy and electronic structure in the gas sensing mechanism.

## 2 Computational details

In the theoretical calculations conducted in this work, the Vienna *ab initio* simulation package (VASP) [35] was employed. The projector augmented wave (PAW) method and the Perdew–Burke–Ernzerhof (PBE) functional within the generalized gradient approximation (GGA) [36,37] were used to describe the electron exchange-correlation interactions. To eliminate interactions between periodic images, a vacuum layer of 15 Å was introduced along the *z*-axis of the BP slab [38]. A plane-wave cutoff energy of 500 eV was adopted, and structural relaxation was performed until the residual forces on each atom were below 0.02 eV/Å. A Monkhorst–Pack *k*-point grid of 3 × 3 × 1 was used for Brillouin zone sampling. The convergence thresholds for electronic self-consistent iterations were set to $10^{-5}$ eV for energy and $10^{-4}$ eV for electronic relaxation, and the van der Waals interactions were described using the DFT–D3 method [39].

The adsorption energies ($E_{ads}$) for gas molecules adsorbed on BP are calculated via Eq. (1):

$$E_{ads} = E_{gas+BP} - E_{gas} - E_{BP} \quad (1)$$

where $E_{gas+BP}$, $E_{gas}$, and $E_{BP}$ represent the energies of gas molecules adsorbed on BP, the isolated gas molecules, and BP, respectively. The charge transfer ($\Delta Q$) is calculated via Eq. (2):

$$\Delta Q = Q_{gas+BP} - Q_{gas} \quad (2)$$

where $Q_{gas+BP}$ represents the charge of gas molecules adsorbed on BP and $Q_{gas}$ represents the charge of the isolated gas molecules. These calculations help to further reveal the underlying mechanism of the gas sensing performance of BP.

In the machine learning section of this study, Python 3.7 was used to implement and evaluate the models. The gas response behavior was formulated as a classification problem, enabling the application of several widely used classification algorithms. Specifically, decision tree (DT) [40], random forest (RF) [41], support vector machine (SVM) [42], K-nearest neighbors (KNN) [43], extreme gradient boosting (XGBoost) [44], and ET [44] were employed. These algorithms were implemented using the scikit-learn library [45]. Before model training, all features were standardized using the StandardScaler method to eliminate differences in units and enhance model stability. All models were optimized through a grid search strategy combined with 5-fold cross-validation for hyperparameter tuning. This systematic approach effectively reduces the risk of overfitting and identifies the optimal parameter set for each model. For model validation, we randomly split the dataset and reserved 20% as an independent test set. During training, we calculated the mean accuracy and standard deviation from cross-validation to assess the robustness and generalization capability of each algorithm [46]. In model training, we primarily relied on a more consistent DFT dataset to ensure comparability and reliability, focusing on the intrinsic relationship between electronic structure descriptors and gas sensing responses rather than reproducing fluctuations caused by sample-dependent differences observed in experiments, thereby ensuring that the model captures fundamental physicochemical trends rather than artifacts of experimental variability.

## 3 Results and discussion

Figure 1(a) shows the top and side views of monolayer BP, which exhibits a characteristic honeycomb structure. Here, 21 gas molecules were selected because they are widely representative in practical industrial and environmental scenarios and are commonly studied in BP gas-sensing research, ensuring strong relevance, comparability, and reference value. The three representative adsorption sites, including hollow (H), top (T), and bridge (B) sites, were chosen according to the structural characteristics of the BP surface (Fig. 1(b)). As shown in Fig. 1(c), the band structure of monolayer BP was calculated using the GGA method. The results indicate that BP is a direct band gap semiconductor, with both the CBM and VBM located at the Γ point in the Brillouin zone, corresponding to energies of −1.91 and 2.76 eV, respectively. The calculated band gap is 0.85 eV, which is slightly smaller than the experimentally observed value. This underestimation is primarily due to the inherent limitations of the GGA method in handling electron–electron interactions, particularly its tendency to underestimate the binding energy of localized states in low-dimensional systems [47]. The PBE exchange-correlation functional was primarily chosen because PBE performs reliably in predicting relative quantities such as adsorption energy, charge transfer, and electronic structure variations. Although more accurate hybrid functional methods such as HSE06 were not employed due to their high computational cost, the GGA approach still effectively captures the essential features of the electronic structure while significantly reducing computational time and resource consumption [48]. This ensures the reliability and comparability of the predicted results. As shown in Fig. 1(d), the density of states (DOS) analysis reveals that the P 2p orbitals contribute most significantly to the total DOS, indicating that they dominate the band structure. Further analysis shows that $\varepsilon_p$ is located at −2.237 eV, suggesting that electrons in the valence band are primarily distributed within the p orbitals. Moreover, the asymmetric distribution of electronic states near the Fermi level supports the p-type conductivity characteristic of BP [49]. Figure 1(e) presents the calculated work function of monolayer BP. Based on the surface electrostatic potential difference, the vacuum energy level is found to be 2.23 eV and the Fermi level is −2.54 eV, resulting in a work function of 4.77 eV. This value is of great importance for further investigations of charge transfer behavior between BP and target gas molecules, as well as for understanding its gas sensing mechanism. Additionally, it provides a theoretical foundation for designing heterojunctions and tuning the interfacial energy band structures in practical applications.

The following analysis examines $NO_2$, a representative gas with a strong response, and $N_2$, a typical nonresponsive gas, to compare the changes in electronic properties after their adsorption on the BP surface. This comparison further reveals how gas adsorption influences the electronic structure and gas sensing performance of the material, thereby providing deeper insights into the underlying sensing mechanism. As a strong oxidizing gas, $NO_2$ can significantly modulate the electronic structure of BP upon adsorption. It can introduce new impurity energy levels, shift band positions, and alter carrier concentration, typically accompanied by pronounced charge transfer and noticeable band structure reconstruction. These changes in electronic structure lead to variations in the material's conductivity, producing a detectable sensing signal. Figure 2 illustrates the electronic structure characteristics of $NO_2$ adsorbed at three typical adsorption sites on the BP surface, including the charge difference density (CDD) distribution, band structure, and DOS. The corresponding atomic structures and work function variations are shown in Fig. S1 in

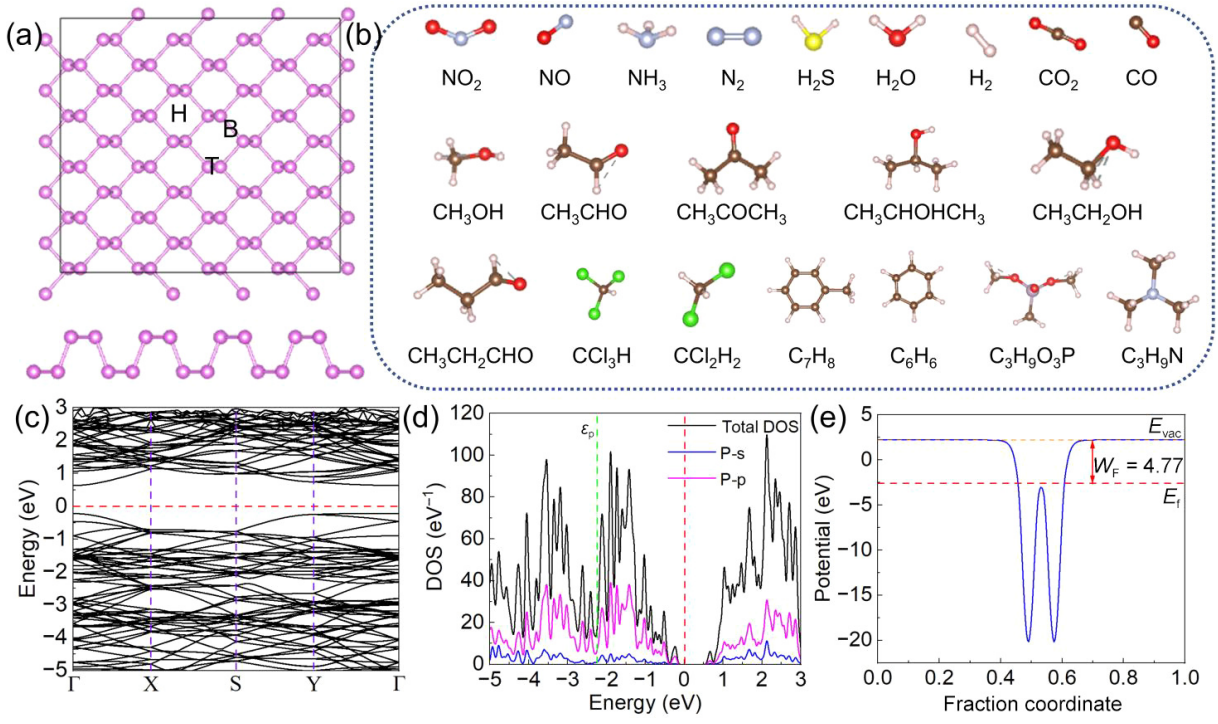



**Fig. 1** Structures of (a) BP and (b) 21 gas molecules, (c) band structures, (d) density of states, and (e) work function of BP. Red dotted line in (c–e) represents Fermi level.

the Electronic Supplementary Material (ESM). Figures 2(a)–2(c) present the CDD distributions of $NO_2$ at the H, T, and B sites, respectively. All three adsorption configurations exhibit evident charge redistribution and interfacial dipole formation. At the H site (Fig. 2(a)), charge accumulation and depletion regions are observed, with an adsorption energy of −0.2097 eV and an adsorption distance of 2.8259 Å. Although interfacial dipoles are formed, the interaction still leans toward physisorption due to the relatively long adsorption distance. In contrast, the T site (Fig. 2(b)) demonstrates an adsorption energy of −0.2395 eV and a charge transfer of 0.2556e, with the shortest adsorption distance of 2.6559 Å, indicating a stronger interaction between the $NO_2$ molecule and surface P atoms. This suggests the possible formation of partial covalent bonds, making this site the most significant in terms of electronic structure reconstruction. The B site (Fig. 2(c)) exhibits an adsorption energy of −0.3012 eV, a charge transfer of 0.2808e, and an adsorption distance of 2.6951 Å, indicating a relatively strong interaction between the $NO_2$ gas molecule and the BP surface. The charge transfer value suggests a significant electronic interaction, while the adsorption distance reflects a balance between the strength of the interaction and the stability of the adsorption site. In this case, the CDD and average electron densities in planes normal all show electron transfer from BP to $NO_2$ (Figs. 2(a)–2(c) and Fig. S2 in the ESM), mainly because $NO_2$ has a high electron affinity and acts as a strong electron-accepting gas, whereas BP possesses a relatively high valence band edge and thus more readily donates electrons. According to the band alignment principle, electrons spontaneously flow from BP to the lowest unoccupied molecular orbital (LUMO) of $NO_2$. Moreover, experimental studies have confirmed that $NO_2$ extracts electrons from BP, leading BP to exhibit a typical p-type sensing response (i.e., a decrease in resistance) [13].

Figures 2(d)–2(f) display the band structure variations of $NO_2$ molecules adsorbed at three different adsorption sites on the BP surface, while Figs. 2(g)–2(i) show the corresponding DOS. As an open-shell paramagnetic molecule with unpaired electrons, $NO_2$ introduces spin polarization effects upon adsorption, leading to distinct spin-up and spin-down band splitting in the electronic structure of BP. Specifically, the adsorption of $NO_2$ introduces impurity states near the Fermi level. These impurity states lie between the CBM and VBM, effectively narrowing the intrinsic band gap of BP. This band gap narrowing can increase the carrier concentration, facilitating electron transitions to the conduction band and thereby enhancing the material's electrical conductivity and gas sensing response. DOS analysis revealed strong coupling between the LUMO of $NO_2$ (mainly 2p orbitals) and the 3p orbitals for the P atoms, where orbital overlap facilitates electron injection into $NO_2$, enhancing the charge transfer and promoting efficient electron injection and extraction, which is critical for gas sensing behavior. This mechanism leads to pronounced electron accumulation beneath $NO_2$ and corresponding electron depletion on the BP surface, representing a typical acceptor-type gas adsorption behavior consistent with the experimentally observed p-type response [13]. Additionally, $NO_2$ adsorption causes the $\varepsilon_p$ of BP to shift closer to the Fermi level, indicating that the valence band states are moving to higher energies. This shift enhances the electron affinity of the material and reflects a stronger ability to attract electrons, thus promoting the charge transfer and improving the sensing sensitivity [50]. From the perspective of work function modulation, $NO_2$ adsorption results in an upward shift of the vacuum energy level and a downward shift of the Fermi level, leading to a significant increase in the overall work function. A higher work function implies that electrons are more tightly bound within the material, making them less likely to escape [51]. This enhances the stability of the adsorption state and strengthens the gas response signal. In short, $NO_2$ adsorption significantly modulates the electronic structure of BP by altering its band structure and DOS, increasing its electron affinity, and changing its work function. These combined effects result in enhanced gas recognition capability and improved sensing performance, highlighting BP's strong potential for detecting $NO_2$ molecules.

In contrast, $N_2$, as a typical inert gas, exhibits highly stable

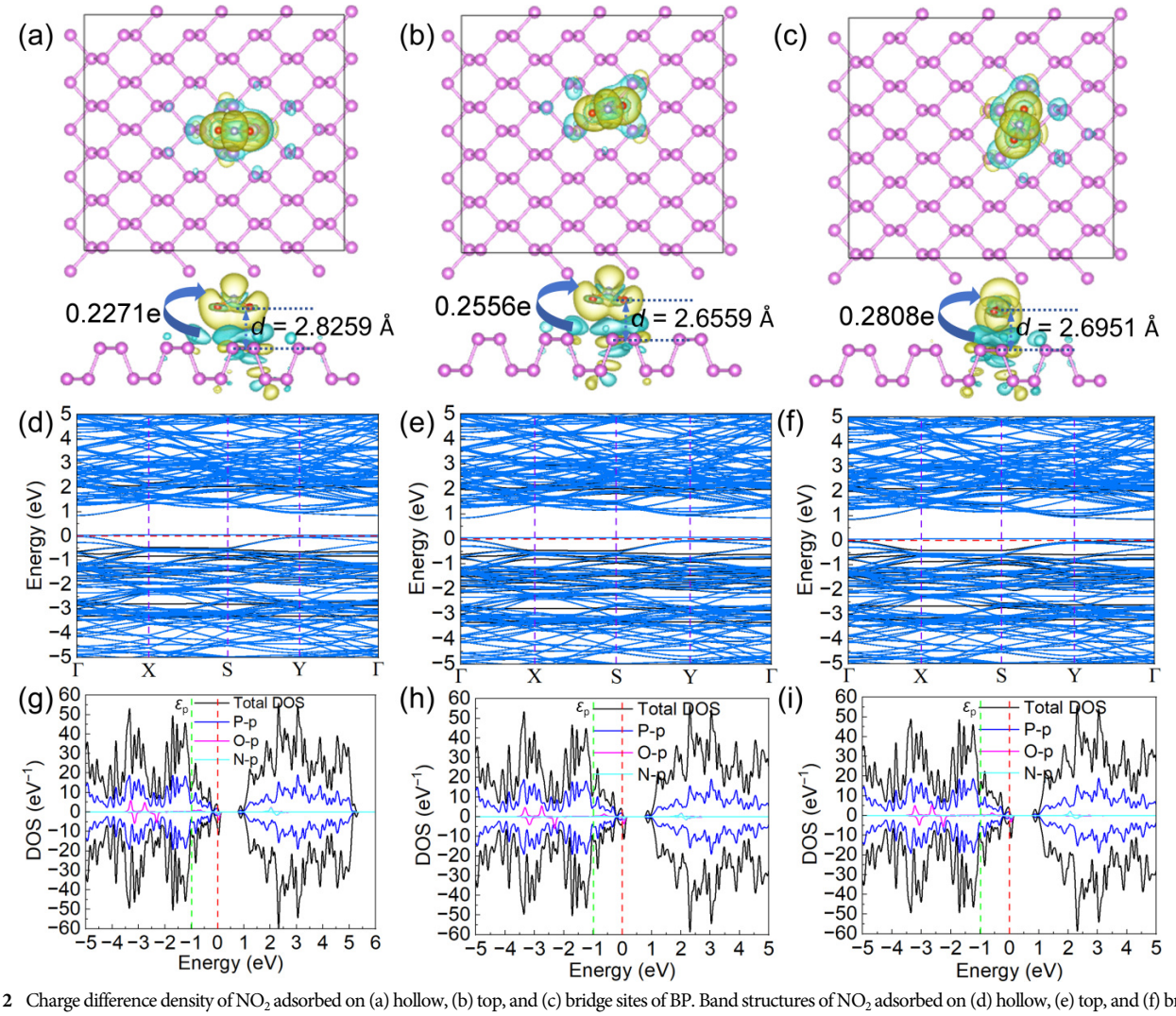



**Fig. 2** Charge difference density of $NO_2$ adsorbed on (a) hollow, (b) top, and (c) bridge sites of BP. Band structures of $NO_2$ adsorbed on (d) hollow, (e) top, and (f) bridge sites of BP. Density of states of $NO_2$ adsorbed on (g) hollow, (h) top, and (i) bridge sites of BP. Yellow and cyan areas in (a–c) represent charge accumulation and depletion, respectively. Red dashed line in (d–i) represents Fermi level, and green dashed line in (g–i) represents p-band center. Black and blue lines in (d–f) represent spin-up and spin-down band structures, respectively.

molecular orbital energy levels and low chemical reactivity, leading to weak interactions with the surface atoms of BP. Regarding adsorption behavior, $N_2$ possesses relatively low $E_{ads}$, indicating that its adsorption state on the BP surface is unstable and dominated primarily by van der Waals forces, characteristic of physical adsorption. Moreover, the adsorption distance between $N_2$ and the surface P atoms is relatively long (Fig. S3 in the ESM), further suggesting a lack of strong bonding with BP, making it difficult to induce structural reconstructions or local electronic state modulations. From the perspective of charge transfer, $N_2$ molecules exhibit minimal electron gain or loss during the adsorption process, with negligible interface charge redistribution (Fig. S3 in the ESM). No significant regions of charge accumulation or depletion are observed, indicating a lack of substantial charge exchange between the gas molecules and BP. Such weak interfacial interactions are insufficient to generate any measurable sensor signal output, resulting in a "nonresponsive" behavior. At the electronic structure level, $N_2$ adsorption introduces virtually no new impurity states or noticeable band perturbations (Figs. S4(a)–S4(c) in the ESM). The DOS distribution remains nearly unchanged before and after adsorption, with no significant shifts in the position of the $\varepsilon_p$ or the emergence of new DOS peaks near the Fermi level (Figs. S4(d)–S4(f) in the ESM). Furthermore, the work function shows no significant change (Figs. S4(g)–S4(i) in the ESM). This suggests that $N_2$ adsorption does not appreciably alter the electronic structure of BP and fails to effectively modulate its conductivity or carrier concentration. In summary, the adsorption of $N_2$ on the BP surface is characterized by weak interactions, low stability, and minimal disturbance to the electronic structure, thereby failing to trigger an effective sensing response, supporting its identification as a nonresponsive gas. This comparison further underscores the importance of active gases, such as $NO_2$, in inducing strong electronic reconstruction and electrical modulation on BP surfaces, providing solid evidence for understanding the gas sensing mechanism. Therefore, the systematic comparison between "strongly responsive" gases (e.g., $NO_2$) and "nonresponsive" gases (e.g., $N_2$) in terms of their structural and electronic properties not only intuitively reveals the structural and electronic synergy in BP under different gas interactions but also offers theoretical guidance for evaluating the responsiveness of other gas species. This analytical approach lays a solid foundation for constructing universal models to predict gas sensing performance and provides valuable theoretical support for designing highly selective and sensitive gas sensors.

To obtain a sufficient and representative dataset for machine

learning, we systematically investigated the gas sensing behavior of BP toward 21 gas molecules by combining a literature review with first-principles calculations. The results reveal that although some adsorption sites are not thermodynamically the most stable, they can still adsorb gas molecules and significantly alter the structural and electronic properties of BP. Therefore, it is both reasonable and necessary to consider the individual characteristics of all possible adsorption sites when constructing predictive models for gas sensing performance, as this approach enhances the model's generalization capability and prediction accuracy. Figure 3 presents the boxplots for the distribution of 9 key physical features derived from the adsorption of 21 different gas molecules on three representative adsorption sites of BP. The figure compares the characteristics between samples with gas responses labeled "0" (nonresponsive) and "1" (responsive), including $E_{ads}$, adsorption distance ($d$), $\Delta Q$, $\varepsilon_p$, VBM, CBM, $E_f$, bandgap ($E_g$), and work function ($W_F$). This comprehensive comparison provides insights into how different physical parameters influence and distinguish the gas sensing behavior. Among these features, the adsorption energy $E_{ads}$ exhibits the most prominent difference between the two response groups, and the heatmap of $E_{ads}$ is shown in Fig. S5 in the ESM. Samples with a response label of "1" generally show lower adsorption energies, indicating stronger interactions and more stable adsorption between the gas molecules and the material. Therefore, $E_{ads}$ is regarded as one of the most representative indicators of gas-sensing activity (Fig. 3(a)). The adsorption distance $d$ also shows a slight decrease in the responsive samples, suggesting that closer atomic contact may facilitate the reconstruction of interfacial electronic states and enhance sensing behavior (Fig. 3(b)). Additionally, $\Delta Q$ tends to be greater in the responsive samples, reflecting more significant electron exchange during the adsorption process (Fig. 3(c)). This enhanced interfacial dipole interaction can contribute to a stronger sensing signal. Although the $\varepsilon_p$ shows a smaller difference between groups, it still reflects a trend of band level regulation related to valence band electronic density (Fig. 3(d)). Regarding band structure features, VBM and CBM exhibit clear trends: VBM shifts upward and CBM moves downward in responsive samples, indicating that gas adsorption induces notable modulation of the electronic band structure (Figs. 3(e) and 3(f)). This modulation is critical for changing carrier concentration or mobility, which is closely linked to sensing performance. $E_f$ shows relatively subtle variations, but they can still describe the overall shift in band positions (Fig. 3(g)). The negligible difference in $E_g$ between responsive and nonresponsive samples indicates that gas adsorption primarily affects sensing behavior through band edge modulation rather than direct alteration of the band gap (Fig. 3(h)). In particular, the variation in $W_F$ in combination with $E_f$ can be used to evaluate the surface potential barrier and electron escape ability, which provides theoretical support for analyzing the interfacial charge transport behavior between the material and external environment (Fig. 3(i)). In summary, the feature distributions reveal the coupling mechanisms between structural reconstruction and electronic modulation during the gas response process. These quantified descriptors not only deepen the understanding of the sensing mechanism but also offer valuable data-driven guidance for the selection of key predictors in machine learning models. With the aid of first-principles calculations, it is possible to rapidly screen and evaluate the sensing capabilities of candidate materials without extensive experimental trials, providing a solid theoretical foundation and a technical pathway for the rational design and optimization of high-performance gas sensing materials.

A further analysis of gas molecules at different adsorption sites (T, H, B) reveals that each site exhibits distinct differences in adsorption strength, structural reconstruction, and electronic structure changes, which directly determine the interaction mode of the gas molecules and the resulting sensing response (Fig. S6 in the ESM). Overall, the H site generally shows a more negative $E_{ads}$, indicating that gas molecules can form a more stable adsorption configuration at this site. Stronger interactions also lead to a smaller distance and more significant charge transfer, suggesting that gas molecules can more easily accept or donate electrons to the material, thereby causing noticeable changes in carrier concentration and conductivity, which is conducive to achieving a higher response intensity. In contrast, the T and B sites exhibit weaker adsorption energies and less charge transfer than the H site, indicating that physical or weak chemical adsorption dominates at these sites, resulting in limited electronic perturbation and generally lower sensing responses. Although

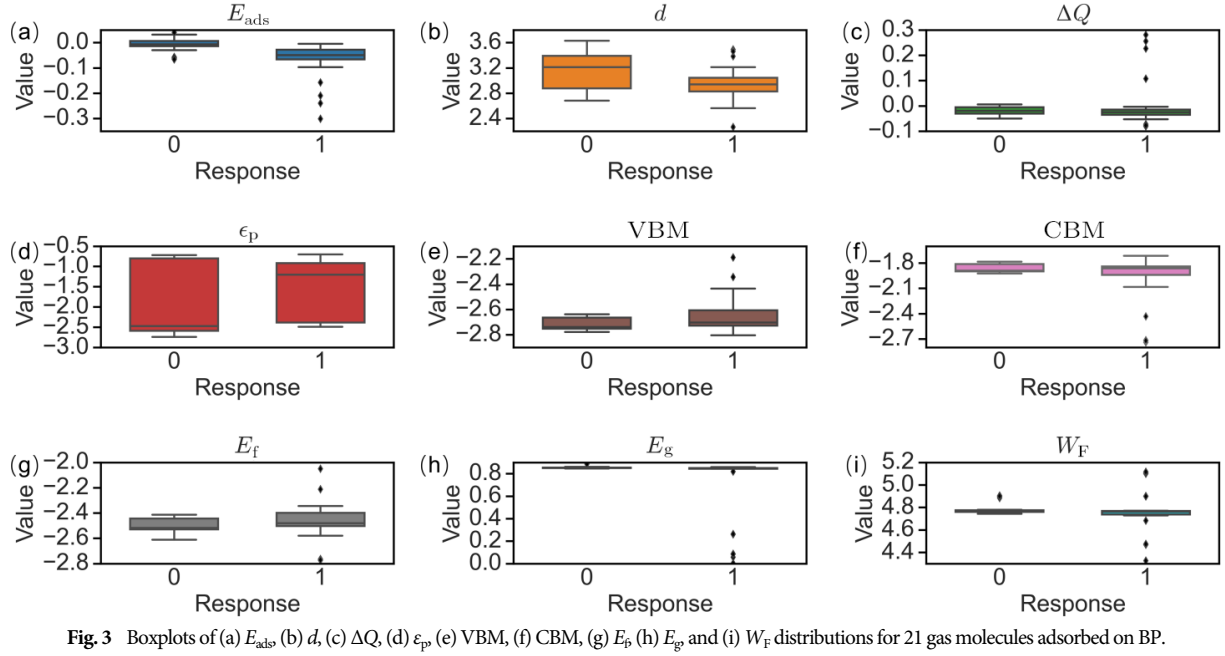



**Fig. 3** Boxplots of (a) $E_{ads}$, (b) $d$, (c) $\Delta Q$, (d) $\varepsilon_p$, (e) VBM, (f) CBM, (g) $E_f$, (h) $E_g$, and (i) $W_F$ distributions for 21 gas molecules adsorbed on BP.

electronic structure parameters such as VBM, CBM, $E_F$, and $E_g$ do not show pronounced differences among the sites, these subtle changes in electronic states can still affect the local density of states and carrier transport pathways, further modulating the sensing behavior. Particularly at the H site, where adsorption is stronger, gas molecule-induced band bending and charge redistribution are more pronounced, making the intrinsic electronic structure of the material more susceptible to perturbation and thereby amplifying the sensing response. In summary, the H site, with its stronger adsorption capability and more significant electronic modulation than the T and B sites, generally exhibits higher sensitivity in the sensing process, whereas the T and B sites, due to weaker interactions with gas molecules, tend to show lower response amplitudes. These findings also indicate that the choice of adsorption site plays a critical role in gas sensing performance.

Based on the gas sensing properties of BP toward 21 different gas molecules, a dataset was constructed containing three typical adsorption sites, resulting in a total of 63 samples. The gas sensing response of BP to each gas molecule was used as the target label, with a response labeled "1" and a nonresponse labeled "0". The numbers of labels "1" and "0" are 33 and 30, respectively, and the label distribution is shown in Fig. S7 in the ESM. During the analysis, the dataset was divided into training and test sets at a ratio of 8 : 2, and all features were normalized prior to modeling. Subsequently, six mainstream machine learning algorithms were employed for model training and comparative analysis, including DT, RF, SVM, KNN, XGB, and ET. The overall modeling process is illustrated in Fig. 4(a). After hyperparameter tuning and optimization, the prediction accuracy and standard deviation of the six algorithms are shown in Fig. 4(b). Among them, the ET model exhibited the best overall performance, achieving a test accuracy of 0.96 with a standard deviation of only 0.08, indicating not only excellent predictive accuracy but also strong robustness and stability across different data partitions. The XGB and RF models also showed relatively high accuracies (0.84 and 0.86, respectively); however, their slightly higher standard deviations suggest some fluctuations in performance across different training subsets. In contrast, the SVM model showed the lowest accuracy (0.68) and the highest standard deviation (0.29), indicating instability and poor generalization ability in this task. Further analysis of the learning curves in Fig. S8 in the ESM reveals that the ET model maintains consistently high training accuracy and a steadily increasing cross-validation accuracy across different training set sizes, demonstrating excellent generalization and learning stability. The RF and XGB models also exhibited reasonable generalization trends. In comparison, the DT model showed signs of overfitting, while the KNN and SVM models suffered from underfitting or weak generalization. In summary, the ET model outperforms the others in terms of accuracy, stability, and robustness, making it the most suitable machine learning algorithm for predicting the gas sensing performance of BP in this study.

Figure 5 systematically evaluates the performance of six commonly used machine learning models in binary classification tasks using the receiver operating characteristic (ROC) curve and the precision-recall (PR) curve, comprehensively reflecting each model's accuracy and robustness in predicting gas sensing responses. As shown in the ROC curves of Fig. 5(a), the ensemble learning models RF, XGB, and ET all achieved perfect area under the curve (AUC) values of 1.00, indicating excellent discriminative

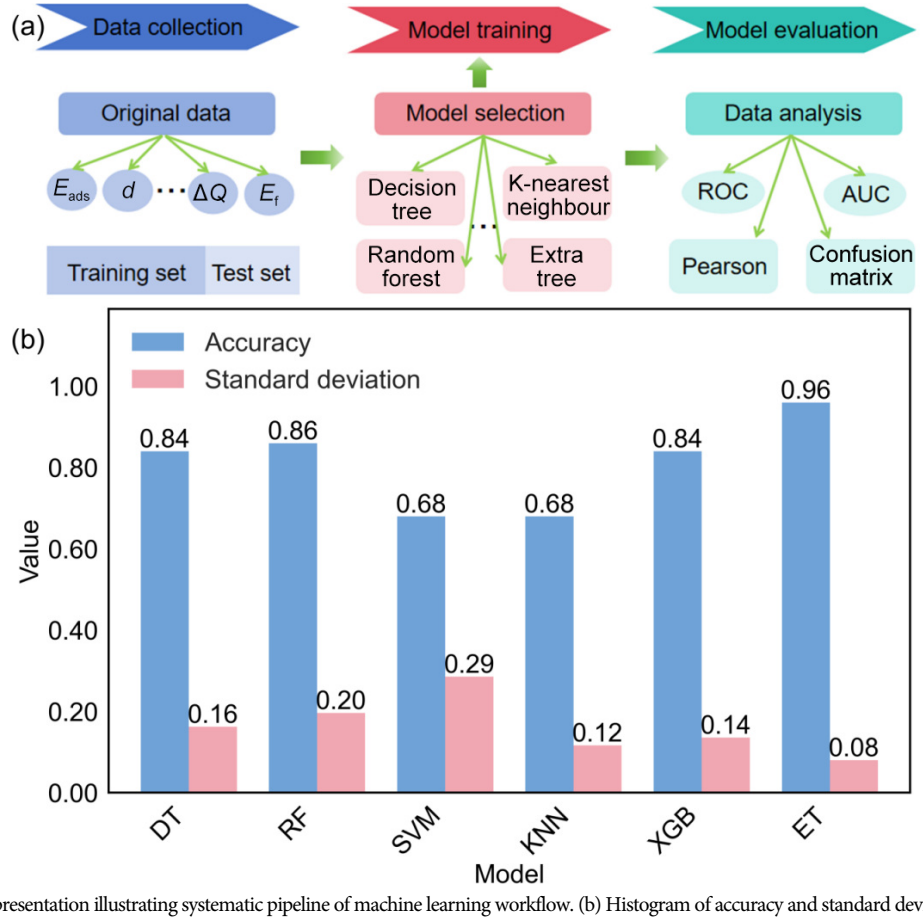



**Fig. 4** (a) Graphical representation illustrating systematic pipeline of machine learning workflow. (b) Histogram of accuracy and standard deviation of data for different models after 5-fold cross-validation.

power and classification precision in distinguishing between response values "0" and "1". The SVM model also demonstrated strong performance with an AUC of 0.95, while KNN and DT achieved AUC values of 0.86 and 0.88, respectively. Although slightly inferior to the ensemble models, they still offer acceptable classification capabilities. The PR curves in Fig. 5(b) further verify the model performance under class imbalance conditions and scenarios requiring high precision. RF, XGB, and ET again achieved excellent performance with average precision (AP) values of 1.00, suggesting that they not only accurately identify positive samples but also maintain extremely low false positive rates, making them well-suited for high-reliability classification tasks. SVM also performed reliably with an AP of 0.94, while KNN (0.77) remained acceptable. In contrast, the DT model has an AP of only 0.71, indicating a weaker balance between precision and recall and a higher risk of misclassification or false negatives. In summary, ensemble models such as ET, RF, and XGB exhibited the best overall performance in this study, offering excellent generalization ability and model stability. These models are particularly suitable for high accuracy and high robustness applications, such as gas adsorption identification and gas sensing response prediction. The SVM algorithm, based on margin optimization, also shows strong competitiveness, especially in small sample or high-dimensional feature spaces. Although KNN and DT models are structurally simple and computationally efficient, they may suffer from performance limitations when handling complex or strongly nonlinear datasets and are therefore more appropriate as baseline models for comparison or lightweight classification tasks.

As indicated by the above analysis, the ET model exhibits the best overall performance. Therefore, the following analysis will primarily focus on the ET model. Figure 6 systematically presents the performance of the constructed ET model on two test sets, evaluated using confusion matrices and four key classification metrics: accuracy, recall, precision, and F1 score. The confusion matrices in Figs. 6(a) and 6(b) illustrate the model's classification results under different test set sizes. As shown in Fig. 6(a), for the 20% test set ($n$ = 13), the model successfully classified all samples with perfect accuracy, and every sample labeled "0" and "1" was correctly identified, resulting in a 100% prediction accuracy. In Fig. 6(b), for the 50% test set ($n$ = 32), the model also achieved excellent classification performance, with only one misclassified sample. The overall accuracy remained as high as 96.9%, demonstrating strong generalization and robustness. Figure 6(c) further quantifies the model's performance through the four key metrics on both test sets. For the 20% test set, the model achieved perfect scores across all metrics: an accuracy, recall, precision, and F1 score of 1.000, indicating not only complete correctness in distinguishing between responsive and nonresponsive gases but also a perfect balance between precision and recall. For the 50% test set, despite the increased sample size, the model continued to perform at a very high level: an accuracy of 0.938, a recall of 0.933, a precision of 0.933 (minimal false positives), and an F1 score of 0.933, reflecting strong harmony between precision and recall. Overall, the ET model demonstrates stable and outstanding classification performance across both test sets. Notably, it maintains high precision and a low false positive rate while ensuring perfect recall. This indicates that the ET model not only fits the training data well but also possesses excellent generalizability when applied to unseen data. The excellent performance of the ET model in this study is primarily attributed to its clear advantages in handling strongly nonlinear, high-dimensional data with complex feature interactions. By introducing greater randomness in both feature selection and split thresholds, ET can explore the feature space more comprehensively, effectively reducing overfitting and enhancing robustness against noise and structural complexity. Therefore, the ET model is particularly suitable for real-world applications that require high classification accuracy, such as identifying gas adsorption behavior and classifying material gas sensing response, offering a reliable foundation for subsequent feature mechanism analysis.

As shown in the Pearson correlation coefficient heatmap in Fig. 7(a) and Fig. S9(a) in the ESM, there are notable correlations among several key features, as well as between these features and the gas sensing response. Specifically, adsorption energy $E_{ads}$ exhibits a strong negative correlation with charge transfer $\Delta Q$ (correlation coefficient of −0.78), indicating that stronger adsorption of gas molecules on the BP surface is typically accompanied by more significant charge transfer. Meanwhile, a positive correlation is observed between adsorption energy and adsorption distance (correlation coefficient of 0.55), suggesting that weaker adsorption (i.e., higher adsorption energy) usually corresponds to a larger gas surface separation. In contrast, the correlation between adsorption distance and charge transfer is relatively weak (−0.21), indicating no obvious linear relationship between them. As shown in the linear fitting results (Figs. S9(b) and S9(c) in the ESM), the relationship between the adsorption energy and adsorption distance can be expressed as $E_{ads} = 0.11d - 0.38$, and that between the adsorption energy and charge transfer can be expressed as $E_{ads} = -0.71\Delta Q - 0.04$. The corresponding $R^2$ values (0.31 and 0.61, Fig. S9(d) in the ESM) indicate that $\Delta Q$ serves as a more reliable predictor of $E_{ads}$ than $d$. Overall, a decrease in $E_{ads}$ (i.e., stronger adsorption) is

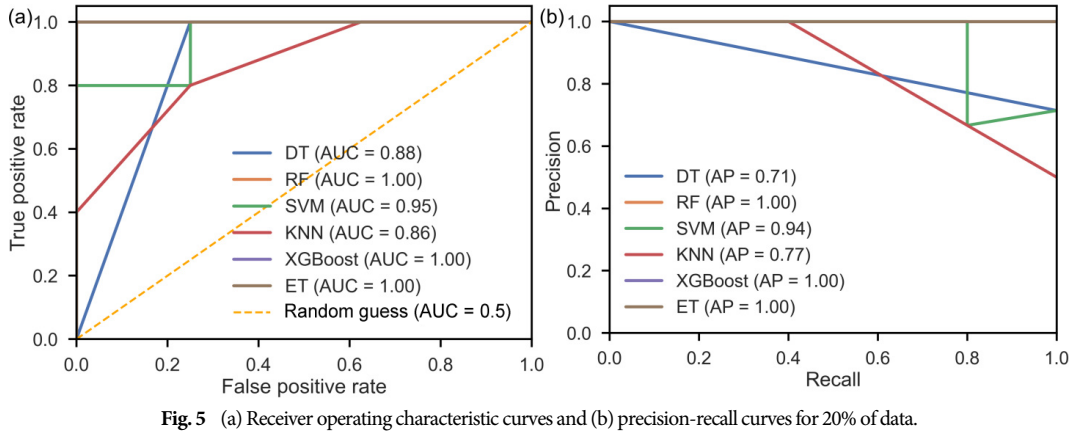



**Fig. 5** (a) Receiver operating characteristic curves and (b) precision-recall curves for 20% of data.

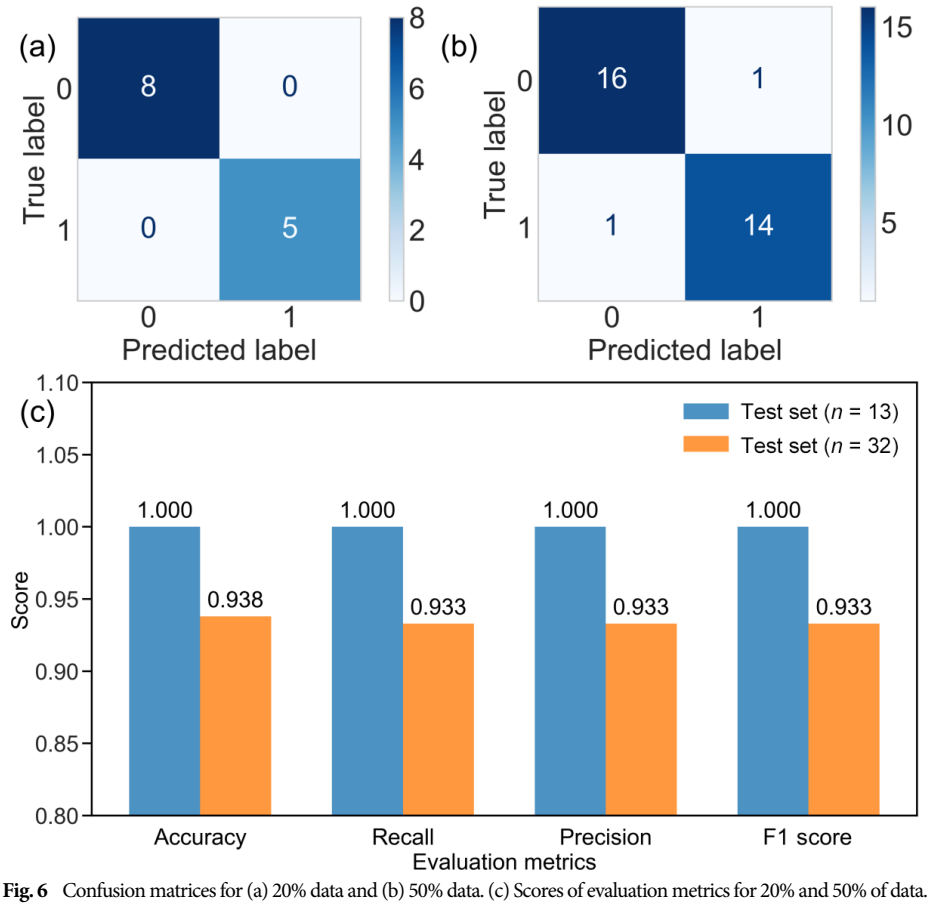



**Fig. 6** Confusion matrices for (a) 20% data and (b) 50% data. (c) Scores of evaluation metrics for 20% and 50% of data.

generally associated with shorter adsorption distances and greater charge transfer. These quantitative correlations clarify the coupling between adsorption geometry and interfacial electron exchange, providing theoretical support for understanding gas adsorption and sensing mechanisms on BP surfaces. In the correlation analysis with the gas sensing response, the adsorption energy (−0.49 eV), adsorption distance (−0.35 Å), CBM (−0.24 eV), and $E_g$ (−0.35 eV) all exhibit relatively strong negative correlations. This indicates that stronger gas adsorption on the BP surface (i.e., lower adsorption energy), shorter adsorption distance, and lower CBM levels are more favorable for triggering gas sensing responses. In contrast, $\varepsilon_p$ (0.32 eV) and VBM (0.24 eV) show positive correlations, suggesting that upward shifts in these energy levels may enhance the electronic response of BP to gas molecules, thereby increasing the likelihood of conductivity changes. In summary, the combined effects of adsorption strength, adsorption distance, and band structure modulation can effectively predict the gas sensing response intensity of different gases on BP surfaces, providing theoretical support for the screening and design of high-performance gas sensing materials.

Figure 7(b) presents a pie chart of feature importance, illustrating the relative contributions of each feature to the model's decision-making process. The results show that $E_{ads}$, CBM, $\varepsilon_p$, VBM, and $E_f$ are the five most representative key features, with contribution rates of 17.2%, 14.4%, 13.4%, 13.2%, and 11.4%, respectively. This indicates that the adsorption strength and local energy level disturbances induced by gas molecules on the BP surface (e.g., $\varepsilon_p$), combined with global band structure variations (VBM, CBM, $E_f$), jointly influence the BP's electronic transport behavior, thus serving as critical criteria for the model to distinguish between "responsive" and "nonresponsive" gases. These findings further confirm the central role of synergistic regulation between adsorption energy and electronic structure in gas sensing mechanisms, highlighting the importance of structure and electronic coupling features in gas sensing performance prediction. Figure 7(c) employs the SHAP method for feature analysis, where red indicates high feature values and blue indicates low feature values. The horizontal axis represents the SHAP value of each feature, reflecting its positive or negative influence on the model's output. Although SHAP differs from traditional feature importance calculation methods, the analysis reveals that $E_{ads}$, $\varepsilon_p$, CBM, VBM, and $E_f$ consistently rank among the top features. This consistency indicates that these five features $E_{ads}$, $\varepsilon_p$, CBM, VBM, and $E_f$ are the most representative and play crucial roles in predicting gas sensing responses. Specifically, $E_{ads}$ directly reflects the strength of gas molecule adsorption on the BP surface and is a key factor governing the sensing response. $\varepsilon_p$ indicates the surface reactivity and the degree of electronic localization; when $\varepsilon_p$ is closer to the Fermi level, the material more readily undergoes electron redistribution upon gas adsorption, enhancing the sensing effect. Since BP is dominated by p-electrons, variations in $\varepsilon_p$ are particularly critical. The VBM and CBM directly influence carrier transitions and conductivity modulation, as gas adsorption can shift the band edges and alter the electrical response. For p-type BP, perturbations near the VBM can significantly affect resistance changes. $E_f$ represents the overall electronic filling state of the system; different gases can induce upward or downward shifts in $E_f$, corresponding to electron donor or acceptor behavior, which directly relates to the oxidizing or reducing properties of the gas species. Therefore, $E_{ads}$, $\varepsilon_p$, CBM, VBM, and $E_f$ are decisive in predicting gas sensing behavior and form a fundamental basis for understanding the sensing mechanism and optimizing material design. Utilizing these five key features enables the development of highly accurate predictive models, offering

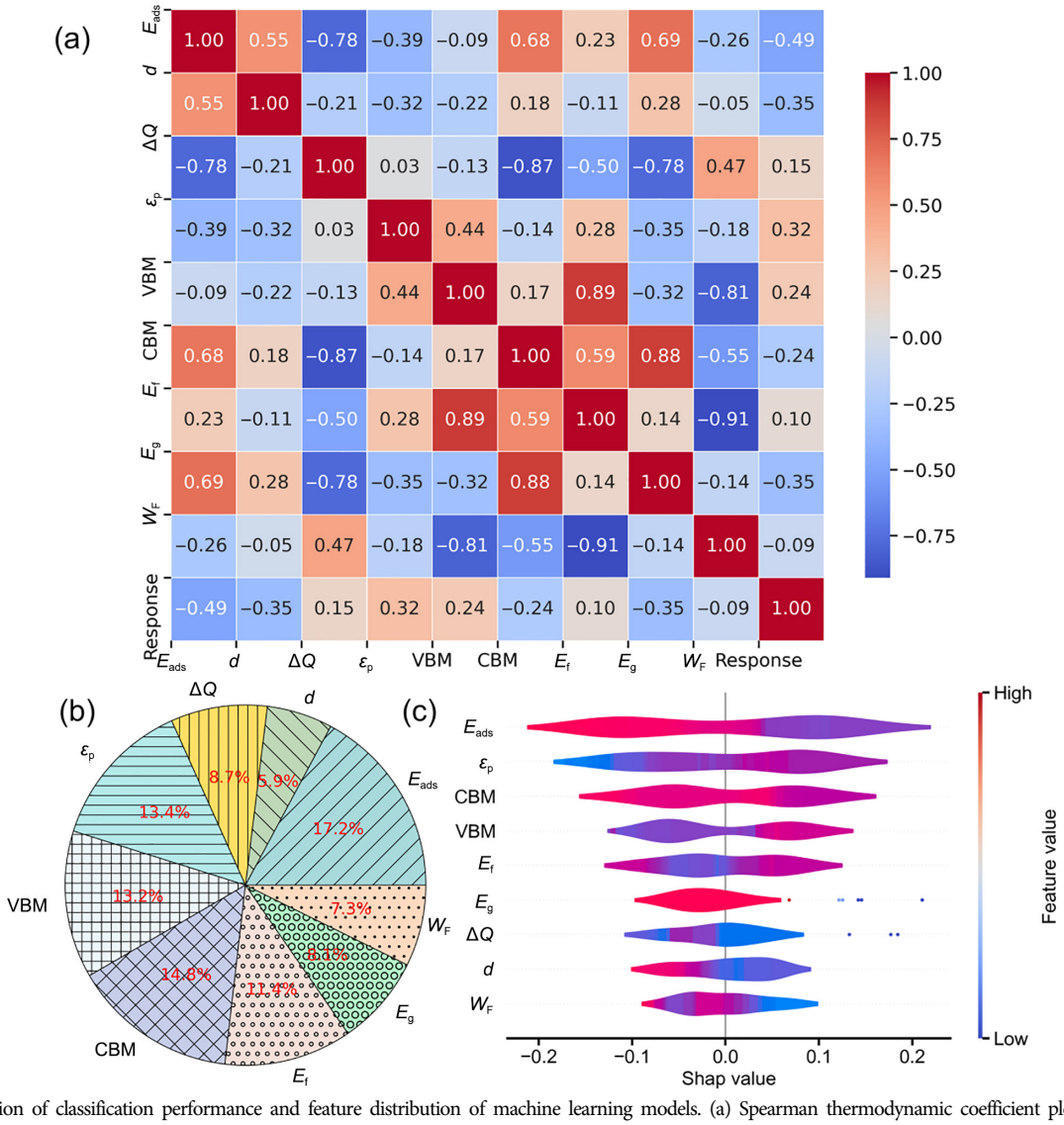



**Fig. 7** Evaluation of classification performance and feature distribution of machine learning models. (a) Spearman thermodynamic coefficient plots of features. (b) Distribution of nine features. (c) Violin plot of SHAP feature importance analysis.

profound insights into electronic structure modulation during gas adsorption. In turn, this study provides a strong theoretical foundation and design strategy for developing gas sensing materials with exceptional selectivity and sensitivity.

Based on the feature importance obtained from the SHAP analysis, the top five most influential features were selected and further analyzed through pairwise combinations to investigate their interactive effects (Fig. 8(a)). The classification performance of each feature pair was evaluated using 5-fold stratified cross-validation (bottom panel). The results show that the $\varepsilon_p$+VBM combination achieved the highest accuracy (0.983), followed by $\varepsilon_p$+CBM (0.953) and VBM+$E_f$ (0.950). Overall, all pairwise combinations of the top five features exhibited high predictive accuracy (all above 0.88), suggesting strong complementarity and representativeness among these features in predicting gas response behavior (Fig. 8(b)). As shown in Fig. 8(c), the decision boundary of the best-performing pair ($\varepsilon_p$+VBM) demonstrates that the model successfully separates response and nonresponse regions, with the lower-left and upper-right areas predominantly classified into different categories. Even though only two features were used, the boundary clearly separates most samples, confirming the strong discriminative capability of this feature pair. In summary, the integration of SHAP-based feature importance analysis and pairwise decision boundary visualization not only reveals the dominant role of key electronic structure parameters in gas response behavior but also validates the significant impact of adsorption energy and band-related features on the model's classification performance, providing quantitative insights for subsequent mechanism analysis and feature optimization.

Next, a lightweight gas response prediction model and visualization system were developed on the Python platform to evaluate the response tendency of BP toward gas molecules. The model was implemented using the Streamlit framework to provide an interactive interface, where five key features determined by SHAP analysis ($E_{ads}$, $\varepsilon_p$, CBM, VBM, and $E_f$) were selected as input parameters. This interface allows researchers to input feature values obtained from first-principles calculations and perform real-time predictions. To enhance the physical interpretability of the model, a feature decision details table was first constructed based on statistical analysis, which includes the specific value of each input feature (Value), its importance in the model (Weight) derived from feature contribution analysis, the feature matching score (Score: 0–1) representing the proximity of the input feature value to the "responsive" sample distribution (normalized to a 0–1 range), and the weighted score (Weighted Score), which represents the overall contribution of each feature

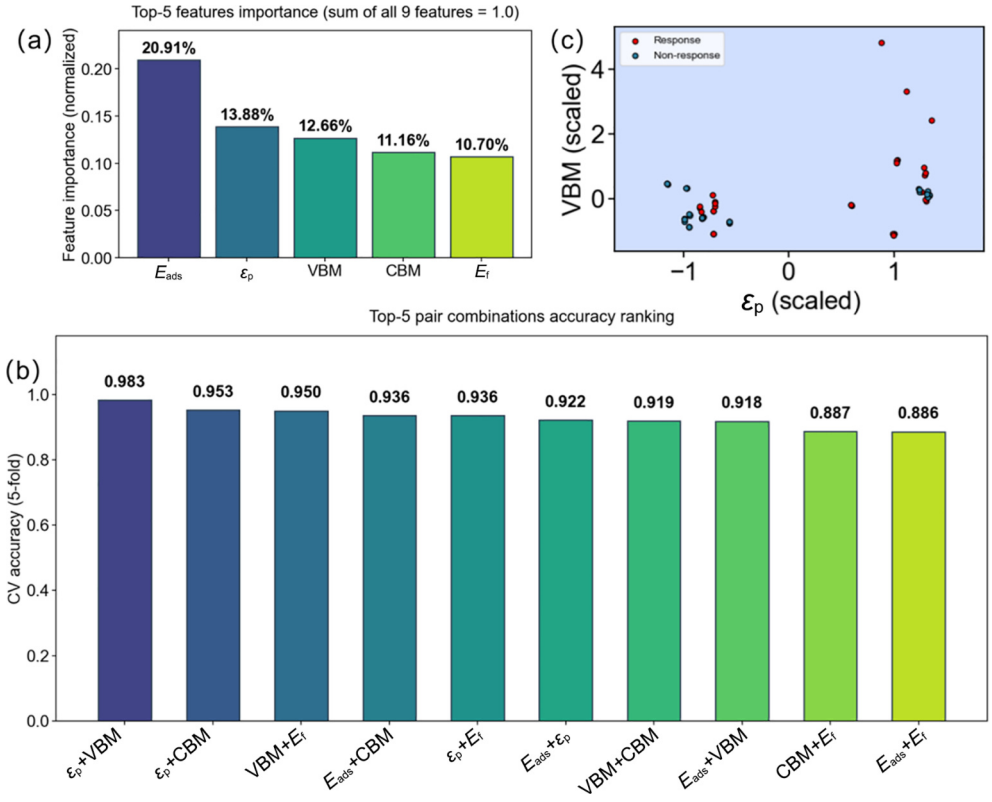



**Fig. 8** (a) SHAP-ranked top five features. (b) Classification accuracy of SHAP-ranked top feature pairs using 5-fold cross-validation. (c) Decision boundary of $\varepsilon_p$+VBM.

after considering its importance (Fig. 9(a)). A higher score indicates that the feature value is closer to the typical range of "responsive" samples, whereas a lower score suggests similarity to "nonresponsive" samples. The overall weighted score, reflecting the integrated influence of all features, is also visualized. For quantitative classification, the threshold value (threshold = 0.241) was determined by analyzing the statistical distribution of the weighted scores and optimizing the ROC curve. This threshold corresponds to the intersection region of the responsive and nonresponsive score distributions, achieving a balanced trade-off between overall accuracy and recall while ensuring reliable distinction of gas response states (Fig. 9(b)). Additionally, a feature value range visualization plot was generated to intuitively illustrate the feature relationships: The green region represents the feature range of responsive samples, the blue region corresponds to nonresponsive samples, and the yellow dot indicates the position of the input sample (Fig. 9(c)). This visualization allows straightforward interpretation of the relative relationship between the input sample and both distributions, thereby supporting a more physically meaningful understanding of the prediction outcome.

Subsequently, the validity of the constructed visualization-based prediction interface was verified. When the five key features ($E_{ads}$, $\varepsilon_p$, CBM, VBM, and $E_f$) obtained from first-principles calculations for the $N_2$ adsorption system on the BP surface were input into the model, the output result was "nonresponse", indicating that BP is insensitive to $N_2$ (Fig. S10 in the ESM). This finding is consistent with previous experimental and theoretical studies, thereby confirming the reliability of the model prediction. Furthermore, to evaluate the generalization ability of the model, ethylene glycol, a molecule not included in the training dataset, was selected as an external test case, and its adsorption feature parameters on the BP surface were calculated (Fig. 9(d)). When these feature values were entered into the model, the system predicted "Response", indicating that BP is sensitive to ethylene glycol molecules. The model's prediction that BP responds to ethylene glycol suggests that the adsorption of ethylene glycol induces significant modulation of the electronic structure of BP. The hydroxyl groups in ethylene glycol can form strong hydrogen bonds or weak coordination interactions with the active sites on the BP surface, leading to electron transfer from ethylene glycol to BP. This electron donation causes shifts in the Fermi level and band-edge positions, triggering a noticeable gas sensing response signal. Such behavior is closely related to the high polarity of ethylene glycol molecules. This result further verifies the model's effectiveness in identifying adsorption systems involving electron-donating molecules and provides theoretical support for exploring the response mechanism of BP toward polar organic gases. Overall, the interpretable prediction framework developed in this study maintains structural simplicity and computational efficiency while providing an intuitive visualization of the contribution of individual electronic structure features to the gas response. It can serve as an effective auxiliary tool for predicting the gas sensing performance of BP and related 2D semiconductors, offering quantitative guidance for experimental parameter optimization, gas selectivity design, and theoretical validation. Moreover, this framework exhibits good generalizability and can be extended to gas adsorption response screening in other two-dimensional material systems, providing new insights for the intelligent design of gas sensing materials. However, the system still has certain limitations. This system relies heavily on the quality and consistency of the input data, and its current predictive capability is primarily based on BP-related datasets. Therefore, cross-material prediction will require more diverse and comprehensive training data. On the other hand, the system provides qualitative trend predictions and is not yet capable of directly predicting quantitative parameters such as response intensity, kinetic behavior, or detection limits.

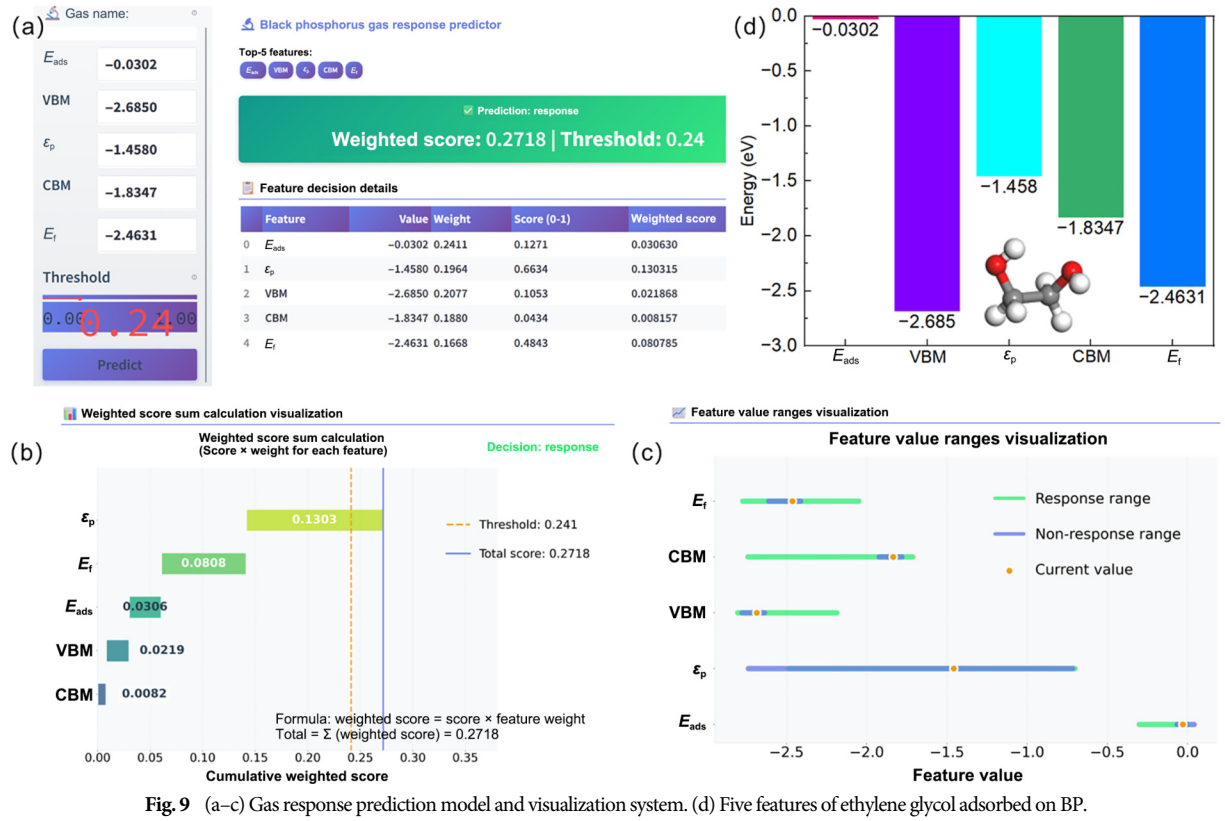



**Fig. 9** (a–c) Gas response prediction model and visualization system. (d) Five features of ethylene glycol adsorbed on BP.

## 4 Conclusions

In summary, to tackle the challenge of accurately predicting whether BP exhibits a sensing response to various gas molecules, this study proposes an intelligent identification framework that integrates first-principles calculations with machine learning techniques. BP was employed as the sensing material, and first-principles calculation simulations were performed to extract key electronic and structural properties at three representative adsorption sites for 21 gas molecules. A suite of machine learning algorithms was subsequently applied to predict the gas response behavior. Among them, the ET model exhibited the best overall performance, with a test accuracy of 0.96 and a standard deviation of only 0.08, and achieved AUC and AP values of 1.00 on both ROC and PR curves, indicating outstanding classification accuracy and robustness. Simultaneously, the ET model achieved high precision predictions on both the 20% and 50% test sets with F1 scores of 1.000 and 0.933, respectively, confirming its strong generalization capability. Moreover, feature importance analysis further revealed the relationship between adsorption energy, adsorption distances, and charge transfer. $E_{ads}$, $\varepsilon_p$, CBM, VBM, and $E_f$ are the most representative descriptors, playing vital roles in determining the sensing response. A lightweight gas response prediction model and visualization system were developed on the Python platform to evaluate the response tendency of BP toward gas molecules. By simply inputting the five key feature parameters obtained from first-principles calculations, real-time prediction of the response behavior of black phosphorus toward different gas molecules can be achieved. Therefore, by combining first-principles calculations with machine learning techniques, the proposed approach introduces an innovative pathway for efficiently screening BP-based gas-sensitive materials while minimizing challenges related to experimental reproducibility. Furthermore, this model holds the potential to be extended to diverse material systems, driving significant advancements in the field of intelligent gas sensors.

### Acknowledgements

This work was supported by the Key Research and Development Plan of Jiangsu Province (No. BE2019094) and the Innovation/Entrepreneurship Program of Jiangsu Province (No. JSSCTD202146). The first author acknowledges the support from the Scientific Research Start-up Funds of Jiangsu University for this work.

### Author contributions

Mingyuan Wang: writing – original draft, visualization, methodology, and investigation. Yaqi Zhang: visualization and formal analysis. Bowen Xiong: visualization and methodology. Ke Wang: visualization and conceptualization. Xiangzhao Zhang: software and conceptualization. Jian Yang: visualization. Lin Xu: writing – review and editing, visualization, and conceptualization. Guanjun Qiao: resources and funding acquisition. Guiwu Liu: writing – review and editing, supervision, resources, and formal analysis.

### Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### Competing interests

The authors have no competing interests to declare that are relevant to the content of this article. The author Guiwu Liu is the Editorial Committee member of this journal.

### Electronic Supplementary Material

Supplementary material (related dataset and the code for

visualization) is available in the online version of this article at https://doi.org/10.26599/JAC.2026.9221243.

## References


[1] Panda S, Mehlawat S, Dhariwal N, *et al.* Comprehensive review on gas sensors: Unveiling recent developments and addressing challenges. *Mater Sci Eng B—Adv* 2024, **308**: 117616.

[2] Khan I, Tariq M, Alabbosh KF, *et al.* Soil microplastics: Impacts on greenhouse gasses emissions, carbon cycling, microbial diversity, and soil characteristics. *Appl Soil Ecol* 2024, **197**: 105343.

[3] Zhang J, Liu XH, Neri G, *et al.* Nanostructured materials for room-temperature gas sensors. *Adv Mater* 2016, **28**: 795–831.

[4] Jang YW, Jo JW, Park SK, *et al.* Room-temperature gas sensors based on low-dimensional nanomaterials. *J Mater Chem C* 2024, **12**: 18609–18627.

[5] Wu KD, Xu ZJ, Xu KC, *et al.* Oxygen vacancy-rich engineering-optimized molybdenum trioxide microbelts for room-temperature ppb-level trimethylamine detection. *J Adv Ceram* 2025, **14**: 9221102.

[6] Zhang NN, Zhou Q, Fan DM, *et al.* Novel roles of hydrocolloids in foods: Inhibition of toxic maillard reaction products formation and attenuation of their harmful effects. *Trends Food Sci Tech* 2021, **111**: 706–715.

[7] Yang XY, Deng Y, Yang HT, *et al.* Functionalization of mesoporous semiconductor metal oxides for gas sensing: Recent advances and emerging challenges. *Adv Sci* 2023, **10**: 2204810.

[8] Li ZH, Zeng W, Zeng FP, *et al.* A highly sensitive ammonia sensor and multifunctional monitoring system based on $MoO_3$@$Ti_3C_2T_X$ (MXene) at room temperature for intelligent agriculture. *J Adv Ceram* 2025, **14**: 9221128.

[9] Zheng W, Lu GC, Liu XH, *et al.* Mixed-dimensional heterojunction by 3D CdS nanowire arrays bridged with 2D $WSe_2$ for ultrafast photoelectric gas sensor. *Matter* 2025, **8**: 101914.

[10] Moseley PT. Progress in the development of semiconducting metal oxide gas sensors: A review. *Meas Sci Technol* 2017, **28**: 082001.

[11] Huang XG, Pang R, Yang MD, *et al.* Flexible gas sensors based on carbon nanotube hybrid films: A review. *Adv Mater Technol* 2023, **8**: 2300616.

[12] Zhao N, Chang X, Liu XH, *et al.* Probing the structural dynamics of $In_2O_3$ using *in situ* Raman spectroscopy: Bridging material dynamics and sensor functionality. *Angew Chem Int Edit* 2025, **64**: e202512808.

[13] Liang ZP, Wang MY, Zhang XZ, *et al.* A 2D–0D–2D sandwich heterostructure toward high-performance room-temperature gas sensing. *ACS Nano* 2024, **18**: 3669–3680.

[14] Du KX, Wang MY, Liang ZP, *et al. In situ* grown large-size highly crystalline black phosphorus thin film for self-driven photodetector. *Appl Phys Lett* 2024, **124**: 063508.

[15] Cho SY, Lee Y, Koh HJ, *et al.* Superior chemical sensing performance of black phosphorus: Comparison with $MoS_2$ and graphene. *Adv Mater* 2016, **28**: 7020–7028.

[16] Zhang DQ, Du Q, Yang L, *et al.* High-throughput experimental technology: Rapid identification of the precious metal modified $In_2O_3$ for $NO_2$ low-temperature sensing. *IEEE Sens J* 2023, **23**: 8101–8108.

[17] Jenkins JL, Kao RYT, Shapiro R. Virtual screening to enrich hit lists from high-throughput screening: A case study on small-molecule inhibitors of angiogenin. *Proteins* 2003, **50**: 81–93.

[18] Zhang R, Chen G, Gao SS, *et al.* Combining first principles and machine learning for rapid assessment response of $WO_3$ based gas sensors. *Int J Min Sci Techno* 2024, **34**: 1765–1772.

[19] Boonpalit K, Kinchagawat J, Prommin C, *et al.* Efficient exploration of transition-metal decorated MXene for carbon monoxide sensing using integrated active learning and density functional theory. *Phys Chem Chem Phys* 2023, **25**: 28657–28668.

[20] Yang ZJ, Sun YJ, Gao SS, *et al.* General model for predicting response of gas-sensitive materials to target gas based on machine learning. *ACS Sensors* 2024, **9**: 2509–2519.

[21] Abbas AN, Liu BL, Chen L, *et al.* Black phosphorus gas sensors. *ACS Nano* 2015, **9**: 5618–5624.

[22] Liu Y, Wang Y, Ikram M, *et al.* Facile synthesis of highly dispersed $Co_3O_4$ nanoparticles on expanded, thin black phosphorus for a ppb-level $NO_x$ gas sensor. *ACS Sensors* 2018, **3**: 1576–1583.

[23] Kou LZ, Frauenheim T, Chen CF. Phosphorene as a superior gas sensor: Selective adsorption and distinct *I–V* response. *J Phys Chem Lett* 2014, **5**: 2675–2681.

[24] Lee G, Kim S, Jung S, *et al.* Suspended black phosphorus nanosheet gas sensors. *Sensor Actuat B—Chem* 2017, **250**: 569–573.

[25] Zhou Y, Hu ZH, Zhao HC, *et al.* Two-dimensional black phosphorus /tin oxide heterojunctions for high-performance chemiresistive $H_2S$ sensing. *Anal Chim Acta* 2023, **1245**: 340825.

[26] Yasaei P, Behranginia A, Foroozan T, *et al.* Stable and selective humidity sensing using stacked black phosphorus flakes. *ACS Nano* 2015, **9**: 9898–9905.

[27] Cho SY, Koh HJ, Yoo HW, *et al.* Tunable chemical sensing performance of black phosphorus by controlled functionalization with noble metals. *Chem Mater* 2017, **29**: 7197–7205.

[28] Donarelli M, Ottaviano L, Giancaterini L, *et al.* Exfoliated black phosphorus gas sensing properties at room temperature. *2D Mater* 2016, **3**: 025002.

[29] Zhao HC, Li J, Wang YJ, *et al.* Black phosphorus nanosheets decorated multiscale zinc ferrite spheres toward swift and humidity-tolerant breath acetone sensing. *IEEE Sens J* 2023, **23**: 1908–1916.

[30] Sarswat PK, Free ML. Real-time detection of thiols using CoPc modified black-phosphorus based sensors. *J Electrochem Soc* 2019, **166**: B1–B8.

[31] Ou PF, Song PF, Liu XY, *et al.* Superior sensing properties of black phosphorus as gas sensors: A case study on the volatile organic compounds. *Adv Theor Simul* 2019, **2**: 1800103.

[32] Zhao HC, Li J, She XP, *et al.* Black phosphorus nanosheets-sensitized Zn-doped α-$Fe_2O_3$ nanoclusters for trace acetone detection. *Sensor Actuat B—Chem* 2023, **395**: 134496.

[33] Liu HH, Shi YB, Wang T. Black phosphorus–carbon nanotube gas sensor for detection of dimethyl methylphosphonate. *Mater Lett* 2021, **293**: 129728.

[34] Korolkov VV, Timokhin IG, Haubrichs R, *et al.* Supramolecular networks stabilise and functionalise black phosphorus. *Nat Commun* 2017, **8**: 1385.

[35] Perdew J, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Phys Rev Lett* 1996, **77**: 3865–3868.

[36] Perdew J, Wang Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys Rev B* 1992, **45**: 13244–13249.

[37] Yang T, Wang D, Zhang Z, *et al.* Elucidating the recombination kinetics of 2-naphthyl radicals with phenylacetylene. *Prog React Kinet Mec* 2025, **50**: 012.

[38] Ahmad W, Aljuhani EH, Alwael H, *et al.* Redox impulse, computational calculation of molecular energy potentials and ultra-trace determination of the food colorant erythrosine b in fruit jams, soft drinks and water. *J Food Compos Anal* 2023, **117**: 105110.

[39] Zhao YL, Wang MY, Liu SW, *et al.* Pt decorated $CoFe_2O_4$/$Co_3O_4$ nanosheets derived from 2D Fe–Co MOF for enhanced HCHO detection. *J Adv Ceram* 2025, **14**: 9221092.

[40] Myles AJ, Feudale RN, Liu Y, *et al.* An introduction to decision tree modeling. *J Chemometr* 2004, **18**: 275–285.

[41] Rigatti SJ. Random forest. *J Insur Med* 2017, **47**: 31–39.

[42] Hearst MA, Dumais ST, Osuna E, *et al.* Support vector machines. *IEEE Intell Syst App* 1998, **13**: 18–28.

[43] Zhang ZH. Introduction to machine learning: K-nearest neighbors. *Ann Transl Med* 2016, **4**: 218.

[44] El Bilali A, Abdeslam T, Ayoub N, *et al.* An interpretable machine

learning approach based on DNN, SVR, Extra Tree, and XGBoost models for predicting daily pan evaporation. *J Environ Manage* 2023, **327**: 116890.

[45] Pedregosa F, Varoquaux G, Gramfort A, *et al*. Scikit-learn: Machine learning in python. *J Mach Learn Res* 2011, **12**: 2825–2830.

[46] Verma BK, Yadav AK. Advancing software vulnerability scoring: A statistical approach with machine learning techniques and GridSearchCV parameter tuning. *SN Comput Sci* 2024, **5**: 595.

[47] Wan ZY, Wang QD, Liu DC, *et al*. Effectively improving the accuracy of PBE functional in calculating the solid band gap via machine learning. *Comp Mater Sci* 2021, **198**: 110699.

[48] Wang HY, Tal A, Bischoff T, *et al*. Accurate and efficient band-gap predictions for metal halide perovskites at finite temperature. *npj Comput Mater* 2022, **8**: 237.

[49] Liu H, Neal AT, Zhu Z, *et al*. Phosphorene: An unexplored 2D semiconductor with a high hole mobility. *ACS Nano* 2014, **8**: 4033–4041.

[50] Li JY, Si WZ, Shi L, *et al*. Essential role of lattice oxygen in hydrogen sensing reaction. *Nat Commun* 2024, **15**: 2998.

[51] Lin L, Jacobs R, Ma TY, *et al*. Work function: Fundamentals, measurement, calculation, engineering, and applications. *Phys Rev Appl* 2023, **19**: 037001.