Materials Science & Engineering R 165 (2025) 101010

Contents lists available at ScienceDirect

Materials Science & Engineering R

journal homepage: www.elsevier.com/locate/mser

A guided review of machine learning in the design and application for pore nanoarchitectonics of carbon materials

Chuang Wang [a], Xingxing Cheng [a,b,*], Kai Hong Luo [c], Krishnaswamy Nandakumar [d], Zhiqiang Wang [a], Meng Ni [e], Xiaotao Bi [f], Jiansheng Zhang [b,c], Chunbo Wang [g]

[a] National Engineering Laboratory for Reducing Emissions from Coal Combustion, School of Energy and Power Engineering, Shandong University, Jinan 250061, China
[b] Shanxi Research Institute of Huairou Laboratory, Taiyuan 030000, China
[c] Department of Energy and Power Engineering, Tsinghua University, Beijing 100084, China
[d] Department of Chemical Engineering, Louisiana State University, Baton Rouge, LA, USA
[e] Department of Building and Real Estate, Research Institute for Sustainable Urban Development (RISUD) & Research Institute for Smart Energy (RISE), The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[f] Department of Chemical Biological Engineering, University of British Columbia, Vancouver, British Columbia, Canada
[g] Department of Energy Power & Mechanical Engineering, North China Electric Power University, Baoding 071003, China

ARTICLE INFO

Keywords:
Porous carbon material
Material design
Machine learning
Energy storage and conversion
Gas adsorption and separation

ABSTRACT

Porous carbon materials have demonstrated significant potential in areas such as carbon capture, gas separation, energy storage, and catalysis, improving energy efficiency and aiding in reducing carbon emissions. With the advancement of global environmental policies, developing efficient and sustainable materials is critical to addressing energy and environmental challenges. However, traditional trial-and-error approaches are often costly and inefficient. Recently, the rapid development of artificial intelligence and machine learning (ML) has introduced data-driven methods to materials science, significantly improving the efficiency of new material development. This review summarizes the application of ML in porous carbon materials, outlining key learning processes and commonly used algorithms, and highlights the latest advancements of ML in porous carbon synthesis and applications, such as carbon capture, energy storage, and supercapacitors. Specifically, it discusses the impact of essential features, such as pore shape, surface area, and pore volume, on different applications, identifies research gaps for non-biomass precursors like coal and tar pitch, and proposes future research directions. This review aims to serve as a resource for ML applications in the field of porous carbon materials, promoting the efficient development and broad application of novel porous materials.

Abbreviation: ANN, Artificial Neural Network; ANFIS, Adaptive Neuro-Fuzzy Inference Systems; BET, Brunauer-Emmett-Teller; CE, Counterfactual Explanations; CNN, Convolutional Neural Network; CMS, Carbon Molecular Sieves; CART, Classification And Regression Trees; DT, Decision Tree; DBN, Deep Belief Networks; DFT, Density Functional Theory; DNN, Deep Neural Network; DBSCAN, Density-Based Spatial Clustering of Applications with Noise; EN, Elastic Network; ESR, Equivalent Series Resistance; FEM, Finite Element Method; GP, Gaussian Processes; GW, Gravitational Waves; GLR, Generalized Linear Regression; GAN, Generative Adversarial Networks; GAP, Gaussian Approximation Potential; GRU, Gated Recurrent Unit; GBT, Gradient Boosting Tree; GBDT, Gradient Boosting Decision Tree; GBRT, Gradient Boosting Regression Tree; IQR, Interquartile Range; KRR, Kernel Ridge Regression; LR, Linear Regression; LOF, Local Outlier Factor; LGB, Light Gradient Boosting; LSTM, Long Short-Term Memory; LGBM, Light Gradient Boosting Machine; LASSO, Least Absolute Shrinkage and Selection Operator; ML, Machine Learning; MD, Molecular Dynamics; MC, Monte Carlo; MAE, Mean Absolute Error; MSE, Mean Squared Error; MLR, Multiple Linear Regression; MLP, Multilayer Perceptron; MSM, Multiscale Modeling; MOF, Metal-Organic Frameworks; MAPE, Mean Absolute Percentage Error; OM, Optical Microscope; OCR, Optical Character Recognition; PV, Pore Volume; PCA, Principal Component Analysis; PDP, Partial Dependence Plots; RT, Regression Tree; RL, Reinforcement Learning; RNN, Recurrent Neural Network; RUL, Remaining Useful Life; RCT, Charge-Transfer Resistance; RMSE, Root Mean Square Error; RMSLE, Root Mean Squared Logarithmic Error; SSA, Specific Surface Area; SVR, Support Vector Regressor; SVM, Support Vector Machine; SEM, Scanning Electron Microscopy; SHAP, Shapley Additive Explanations; TCN, Temporal Convolutional Network; VAE, Variational Autoencoder; ViT, Vision Transformer; XGB, Extreme Gradient Boosting; XPS, X-ray Photoelectron Spectroscopy; XGBR, Extreme Gradient Boosting RegressorEnglish letter; V_{mic}, Micropore Volume; V_{mes}, Mesopore Volume; S_{mic}, Micropore Surface Area; S_{mes}, Mesoporous Surface Area; R^2, R-Square.

* Corresponding author at: National Engineering Laboratory for Reducing Emissions from Coal Combustion, School of Energy and Power Engineering, Shandong University, Jinan 250061, China.
E-mail address: xcheng@sdu.edu.cn (X. Cheng).

https://doi.org/10.1016/j.mser.2025.101010
Received 10 March 2025; Received in revised form 22 April 2025; Accepted 28 April 2025
Available online 3 May 2025

1. Introduction

Over the past several decades, with growing global emphasis on environmental protection and energy demands, the development of efficient, sustainable materials has become a core focus of scientific research [1]. As shown in Fig. 1, carbon materials, encompassing a wide range of types such as activated carbon, activated semicoke, carbon nanotubes, carbon aerogels, fullerenes, graphene, and carbon fibers, exhibit diverse structures and exceptional properties for adsorption, energy storage, and catalysis, making them valuable in chemical engineering, healthcare, environmental protection, and other fields [2,3]. As a research area with a rich history of continuous innovation, carbon materials have attracted increasing attention for their outstanding performance in applications such as energy storage in batteries [4], catalyst supports for hydrogen production via water electrolysis [5], and high-capacity adsorbents for hydrogen and CO_2 storage [6,7].

Industrial production of carbon materials typically requires testing specific raw materials and process parameters to establish a complete process system. Whenever raw materials or product requirements change, parameters must be adjusted and re-evaluated [8]. Due to the varying process pathways and site-specific parameters, there is currently no standardized parameter system available, with most relying on engineering expertise and limited qualitative principles. The structural and compositional diversity of porous carbon materials is vast, making exhaustive experimental exploration nearly impossible [9]. Additionally, the characteristics of porous carbon materials often have complex, nonlinear relationships with desired performance attributes, making structural tuning highly dependent on accumulated experimentation and experience. This iterative trial-and-error process represents a significant technological barrier and growth bottleneck in the industry [10,11]. Consequently, developing more efficient approaches to circumvent redundant and inefficient experimental processes for porous carbon materials is of critical importance.

In any field, data serves as a potential source for knowledge discovery. In recent years, the development of artificial intelligence and ML has bolstered expectations that data-driven materials science could fundamentally transform scientific discovery [12]. For ML, data is evidence of facts, that is, instances or examples of variables or parameters [13]. In the field of carbon materials, decades of experimental accumulation have generated a wealth of multi-dimensional, comprehensive data that create favorable conditions for the application of ML algorithms.

The advancements in ML have introduced a paradigm revolution in materials science research. ML techniques offer a data-driven approach that identifies critical features in the design and application of porous carbon materials by learning from vast amounts of existing data, enabling predictive modeling [14]. ML leverages extensive training data and suitable algorithms to simulate and deduce patterns in physical or chemical processes, making it an ideal alternative to complex theoretical and empirical computations [15]. As illustrated in Fig. 2, with the rapid advancement of computational power, the application of ML in porous carbon material research has evolved from basic performance prediction to more complex tasks such as design optimization, generative modeling, and structural feature recognition [16,17]. ML models have progressed from traditional approaches that handle small datasets to deep learning models that can explore the intricate relationships between material structure and performance [18,19]. In the future, models are expected to evolve towards adaptive, interpretable AI and reinforcement learning, further enhancing the intelligent design, optimization, and application efficiency of porous carbon materials [20].

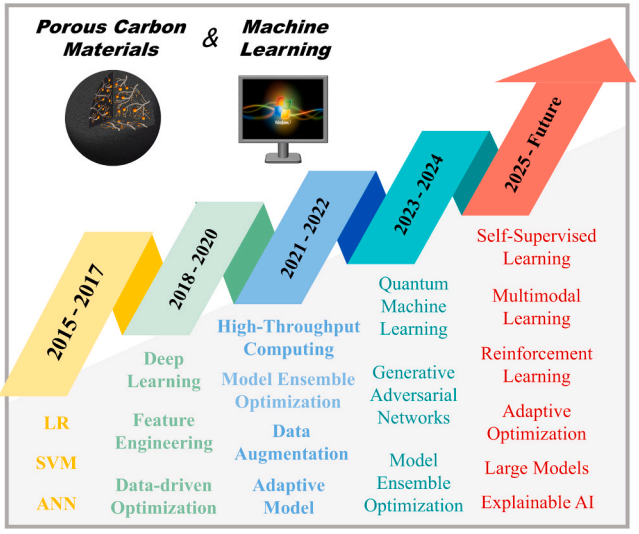


Fig. 2. Keywords for the application of ML in porous carbon materials in recent years.

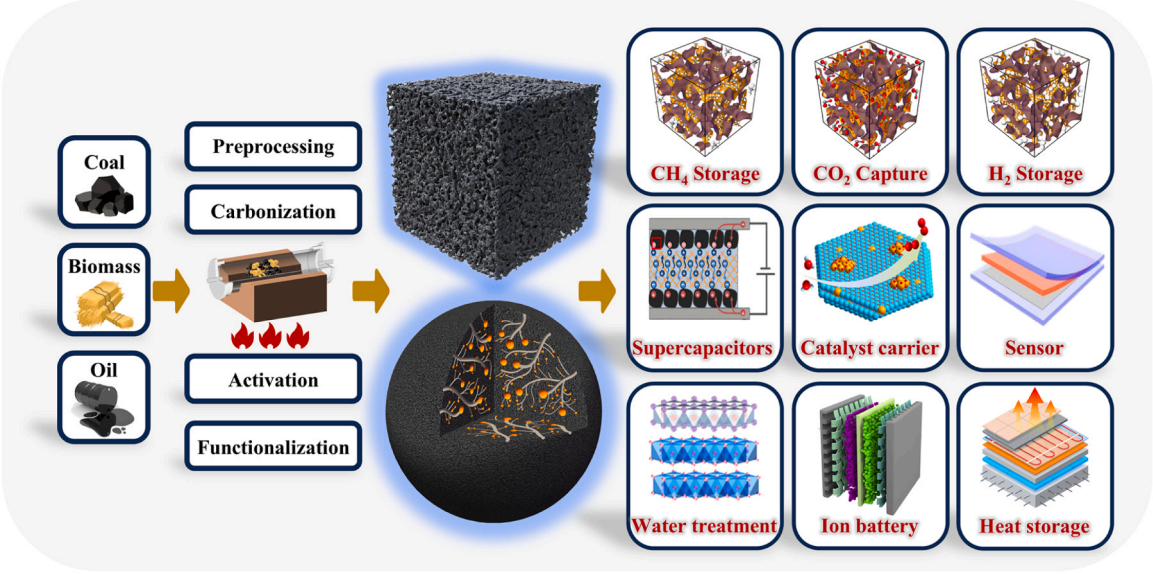


Fig. 1. Wide application of porous carbon-based materials.

In this review, we discuss recent progress in ML-assisted synthesis and application of porous carbon materials. We outline the steps and considerations for applying ML in porous carbon materials and introduce ML methods specific to this field. Finally, we examine the current state and potential opportunities for ML-assisted carbon synthesis and applications, offering insights for future research directions. As illustrated in Fig. 3, this study provides a systematic overview of the research content.

2. Data acquisition

ML algorithms reveal patterns in data through statistical models and optimization techniques. Depending on the chosen algorithm, ML can provide insights through prediction or classification. ML relies on data to train models, enabling the discovery of patterns and making predictions or decisions. Therefore, the quality, quantity, and characteristics of the data directly affect the performance and accuracy of the models [21]. Acquiring high-quality and sufficient materials data has become a core challenge in applying ML in carbon materials science. ML in carbon materials science has somewhat shifted toward addressing the challenges of data acquisition in this field [22].

2.1. Acquisition from literature

Literature data primarily includes experimental and simulation data, typically provided in various formats, with significant differences in quality and consistency [23]. Before transforming this data into a training set, it needs to be filtered and processed, condensing the data into low-dimensional feature descriptors.

2.1.1. Experimental data

The experimental data reported in the literature on porous carbon materials primarily includes pore structure characteristics (e.g., SSA, PV, pore size distribution), physicochemical properties (e.g., carbon content, density, electrical conductivity), adsorption performance (e.g., gas adsorption capacity, adsorption heat), electrochemical performance (e.g., specific capacitance, energy density), structural characterization data (e.g., XRD, Raman spectra, SEM/TEM image features), preparation conditions (e.g., carbon source type, carbonization temperature, type of activating agent), and application performance (e.g., catalytic and energy storage properties) [24]. These datasets can be utilized to meet various machine learning requirements, while image data can also be leveraged through deep learning to extract morphological features of pores, providing valuable support for performance prediction and material optimization.

Models usually perform best when trained on data with a wide range of attributes, necessitating the inclusion of both favorable and less favorable attribute data [25]. In the case of porous carbon materials, literature often presents a larger amount of high-quality data, making it particularly important to supplement comparative data within the materials. ML literature in the same research direction is also an important way to obtain the required data because they have already summarized some literature data. Moreover, when the data contains many missing or low-quality entries, additional experiments may be required to enhance the dataset.

In the literature on the design of porous carbon materials, each study typically contains an average of 3–6 data points, and the number of relevant papers on porous carbon design in the Web of Science has reached thousands, resulting in a substantial data volume. As shown in Table 1, some of the carbon material design data in the literature includes textual data, ratio data, Boolean data, and numerical data of varying magnitudes. The correlation between each feature and the predicted outcome varies, and optimizing the selection of input features (for example, by creating new features from the ratios of two low-correlation features) can significantly impact the overall performance of the model. Some authors may overlook experimental conditions they deem unimportant when writing their papers, leading to missing data. In such cases, missing data can be handled by either discarding incomplete datasets or imputing missing values using default or average values. Similarly, the literature data regarding the applications of porous carbon materials exhibits similar characteristics. Data collection can be done manually or through web crawler techniques [26]. Additionally, models

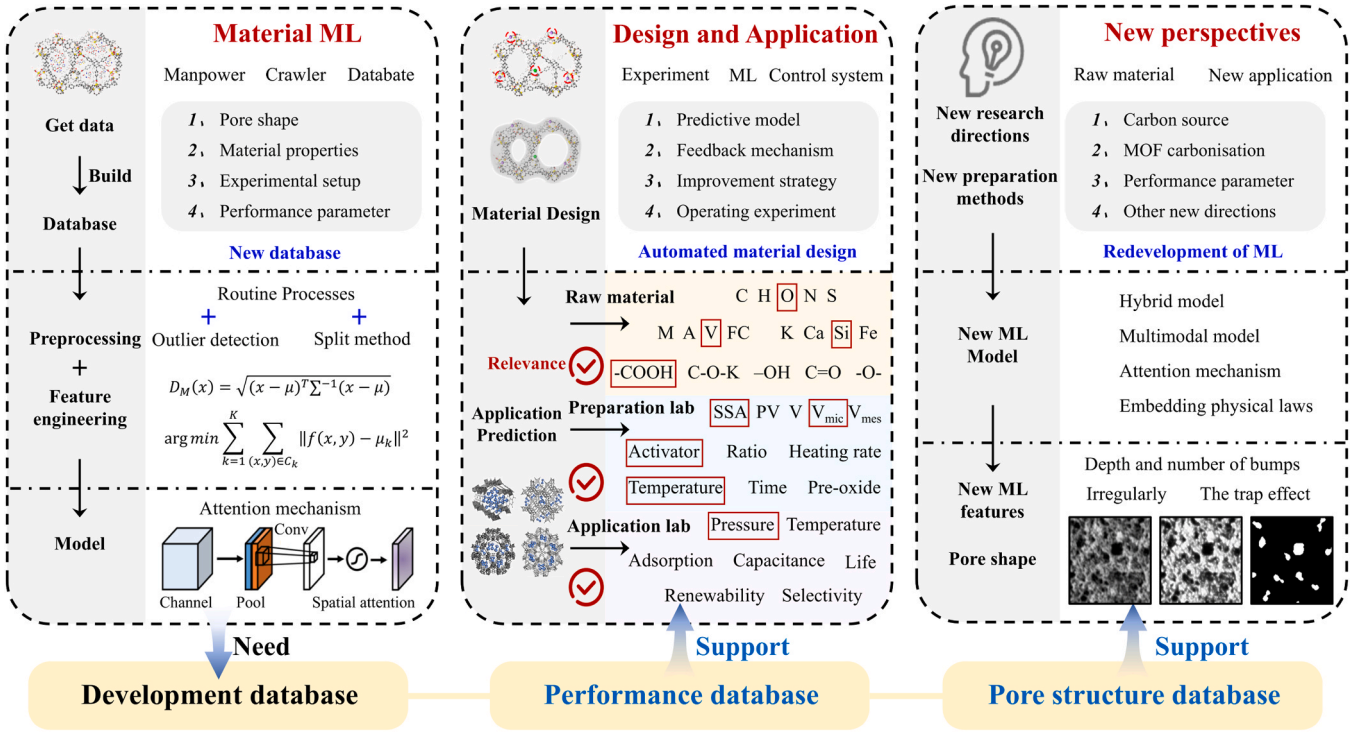


Fig. 3. Overview of the contents of the full paper.

Table 1
Data presentation for carbon material design.

Raw Material	Pre-oxidation	Acid Treatment	Additives	Quality Ratio	Heating Rate ℃/min	Heat Treatment ℃-h	SSA m^2/g	PV cm^3/g	PV (Mic:Mes:Mac)	Refer
Coke	0	0	KOH	3:1	10	750–2	807.7	0.370	0.17:0.62:0.21	[27]
Coke	0	0	KOH	3:1	10	850–2	69.1	0.049	0.18:0.41:0.41	
Coke	0	0	KOH	3:1	10	950–2	113.9	0.155	0.26:0.49:0.25	
Zhundong Coal	0	0	H_3PO_4	2:1	-	950–1	2308	1.510	0.52:0.48:0.00	[28]
Zhundong Coal	0	0	H_3PO_4	2:1	-	500–2	975	0.530	0.95:0.05:0.00	
Zhundong Coal	0	0	H_3PO_4	2:1	-	950–1	1157	0.620	0.90:0.10:0.00	
Liquefied Asphalt	0	0	KOH	2:1	5	600–2	1528	0.660	0.88:0.12:0.00	[29]
Liquefied Asphalt	0	0	KOH	2:3	5	700–2	1937	0.840	0.89:0.11:0.00	
Liquefied Asphalt	0	0	KOH	2:5	5	800–2	2294	1.070	0.85:0.15:0.00	
Liquefied Asphalt	0	0	KOH	2:7	5	900–2	2188	1.170	0.58:0.42:0.00	
Coal Tar Pitch	1	0	H_3PO_4	1:5	4	900–2	1606.6	0.810	0.28:0.60:0.12	[30]
Zhundong Coal	0	1	$Fe_2(C_2O_4)_3$	1:7	8	850–1	449.7	0.354	0.61:0.39:0.00	[31]
Zhundong Coal	0	1	$C_{13}H_{11}ClFeO$	1:10	8	850–1	131	0.333	0.45:0.55:0.00	
Zhundong Coal	0	1	Fe_2O_3	2:5	8	850–1	413.7	0.347	0.58:0.42:0.00	
Zhundong Coal	0	1	$Fe(NO_3)_3$	3:50	8	850–1	436.1	0.327	0.65:0.35:0.00	
Zhundong Coal	0	1	$FeCl_3$	1:10	8	850–1	595.3	0.504	0.56:0.44:0.00	
Subbituminous Coal	1	0	KOH	1:1	8	800–1	939	0.480	0.77:0.23:0.00	[32]
Subbituminous Coal	1	0	KOH	2:1	9	800–1	1021	0.520	0.77:0.23:0.00	
Subbituminous Coal	1	0	KOH	2:1	10	900–1	738	0.350	0.63:0.37:0.00	
Lignite	0	0	$CO(NH_2)_2$	1:1		700–2	459	0.171	0.97:0.02:0.01	[33]
Lignite	0	0	$CO(NH_2)_2$	2:1		700–2	525	0.193	0.97:0.03:0.00	
Lignite	0	0	$CO(NH_2)_2$	3:1		700–2	616	0.228	0.95:0.03:0.02	
Lignite	0	0	$CO(NH_2)_2$	5:1		700–2	460	0.168	0.96:0.04:0.00	
High Sulfur Coal	0	1	KOH	2:1	5	800–2	1864	0.821	0.94:0.04:0.02	[34]
High Sulfur Coal	1	1	KOH	3:1	5	800–2	3345	1.715	0.84:0.13:0.03	
Anthracite	0	0	KOH	3:1	2	700–2	2845.4	1.163	0.94:0.05:0.01	[35]
Anthracite	0	0	KOH	4:1	2	700–2	3550.7	2.168	0.64:0.35:0.01	
Anthracite	0	0	KOH	5:1	2	700–2	3359.7	2.053	0.47:0.53:0.00	
Anthracite	0	0	KOH	6:1	2	700–2	3170.7	1.753	0.39:0.60:0.01	
Bituminous Coal	0	1	KOH	6:1	5	750–1	3160	2.010	0.59:0.28:0.13	[36]
Bituminous Coal	0	1	KOH	6:7	5	750–1	3647	2.570	0.44:0.32:0.24	
Vaporized Slag	0	0	KOH	3:2	10	700–1	747	0.510	0.47:0.53:0.00	[37]
Vaporized Slag	0	0	KOH	2:1	10	700–1	963	0.540	0.59:0.41:0.00	
Vaporized Slag	0	0	KOH	5:2	10	700–1	1231	0.710	0.56:0.44:0.00	
Vaporized Slag	0	0	KOH	3:1	10	700–1	1600	0.900	0.57:0.43:0.00	
Waste Coal	0	0	KOH	1:1	10	400–2	143.8	0.220	0.29:0.66:0.05	[38]
Waste Coal	0	0	KOH	5:2	10	400–2	161.5	0.280	0.27:0.68:0.05	
Waste Coal	0	0	KOH	5:1	10	400–2	258	0.440	0.18:0.76:0.06	
Waste Coal	0	0	KOH	1:1	10	600–2	275.2	0.360	0.66:0.33:0.01	
Graphene	1	1	KOH	2:1	-	800–1	1597.6	0.790	0.75:0.25:0.00	[39]
Graphene	1	1	KOH	4:1	-	800–1	1847.2	0.800	0.69:0.31:0.00	
Graphene	1	1	KOH	6:1	-	800–1	2144.6	1.750	0.47:0.53:0.00	
Graphene	1	1	KOH	8:1	-	800–1	2428.6	1.820	0.37:0.63:0.00	
Graphene	1	1	KOH	10:1	-	800–1	2352.9	1.800	0.41:0.59:0.00	
Industrial Peat	0	0	$KHCO_3$	2:1	10	700–1	745	0.438	0.48:0.44:0.08	[40]
Industrial Peat	0	0	$KHCO_3$	2:3	10	800–1	1135	0.504	0.79:0.20:0.01	
Industrial Peat	0	0	$KHCO_3$	2:5	10	850–1	1325	0.608	0.77:0.22:0.01	
Blue Focus	0	0	K_2CO_3	1:4	10	600–3	238	0.264	0.78:0.22:0.00	[41]
Blue Focus	0	0	K_2CO_3	1:2	10	600–3	302	0.331	0.82:0.18:0.00	
Blue Focus	0	0	K_2CO_3	3:4	10	600–3	587	0.322	0.70:0.30:0.00	
Blue Focus	0	0	K_2CO_3	1:1	10	600–3	812	0.473	0.55:0.40:0.05	
Charcoal Residue	0	1	KOH	1:1	8	800–1	1114	0.366	0.36:0.64:0.00	[42]
Charcoal Residue	0	1	KOH	2:1	9	800–1	1246	0.369	0.42:0.58:0.00	
Charcoal Residue	0	1	KOH	3:1	10	800–1	1377	0.402	0.43:0.57:0.00	
Charcoal Residue	0	1	KOH	3:4	11	800–1.5	1596	1.297	0.35:0.65:0.00	

can employ automated learning and self-updating algorithms, allowing them to continually accumulate data during use, thus achieving an optimized model with learning capabilities.

2.1.2. Simulated data

The simulation data in the literature on porous carbon materials typically includes a variety of aspects, such as pore structure simulations (e.g., pore size distribution, SSA, and PV), adsorption performance simulations (e.g., gas adsorption isotherms, adsorption energy, and adsorption site distribution), electrochemical performance simulations (e.g., charge distribution in electrode materials, ion diffusion behavior, and specific capacitance), thermodynamic and kinetic simulations (e.g., pyrolysis pathways, reaction energy barriers, and kinetic constants during carbonization and activation processes), optoelectronic performance simulations (e.g., light absorption efficiency, electronic bandgap, and carrier mobility), and mechanical performance simulations (e.g., the impact of porosity on structural strength and compression behavior) [43]. These simulation data provide theoretical guidance for structural design, performance prediction, and process optimization of materials. When combined with experimental data, they can be used to construct machine learning models, thereby further improving the efficiency of material design.

2.1.3. Crawler technology

Using web crawling techniques to acquire experimental or simulation data from literature on porous carbon materials enables the rapid collection of data such as specific surface area, pore volume, adsorption performance, electrochemical performance, preparation conditions, and simulation parameters. These data can then be utilized for constructing machine learning models or material databases, as illustrated in Fig. 4. The primary advantage of this approach lies in its ability to efficiently extract large-scale literature data, uncover potential correlations, reduce manual data processing time, and enhance the scope and depth of research through data integration. However, this method also has drawbacks. Crawled data may exhibit inconsistencies or noise, such as difficulties in data standardization due to varying experimental conditions, or the complexity of literature content leading to the omission or misinterpretation of key data [44]. Certain data presented in graphical formats within the literature, such as pore size distribution from BET analyses, specific values from adsorption isotherms, detailed features of electrochemical cyclic voltammetry curves, or Raman spectral parameters (e.g., peak width and intensity ratios), are often challenging to extract directly through crawling. These data typically require further processing using image analysis techniques. Additionally, critical experimental details (e.g., specific instrument parameters and environmental conditions), key simulation parameters (e.g., computational settings and choice of potential functions), and manually derived analyses (e.g., microscopic structural descriptions based on image observations) are difficult to retrieve directly through web crawling. The absence of such data can lead to incomplete datasets, thereby impacting the performance of machine learning models. To address this, supplementary methods such as optical character recognition (OCR) [45] and image analysis tools [46] can be used to extract data from charts and graphs. Furthermore, manual screening and in-depth reading of literature are essential for capturing critical experimental details, thereby improving data completeness and accuracy. It is important to note that web crawling activities may involve copyright issues. Unauthorized crawling could violate legal regulations or the terms of use of journals. To fully leverage the advantages of web crawling while ensuring compliance, a standardized workflow should be established, including data cleaning, normalization, and validation processes. Operations must remain within the bounds of legal frameworks, and expert knowledge should be incorporated to interpret the data appropriately. This approach ensures data quality and enhances the reliability of research outcomes.

2.2. Acquisition from databases

Data in the database undergoes rigorous screening, ensuring high quality and consistency. Data sharing between different research groups is crucial for predicting the structure and performance of porous carbon-based materials [47]. Table 2 lists some permanent storage databases and data collection plans related to carbon materials, providing valuable support for the application of ML in the synthesis and performance prediction of carbon materials. The data in these databases primarily originates from experiments and DFT simulations, while high-throughput computing or experimental techniques are also key methods for generating and expanding datasets [48].

Experimental databases typically include structural information about materials, such as pore size distribution, SSA, adsorption isotherms, and PV, derived from experimental measurements and characterizations. In contrast, simulation databases generate data using methods such as MD and DFT, supplementing data that is difficult to obtain through experiments [49]. Crystal structure databases can also provide detailed information at the molecular and subatomic levels. These databases generally utilize structured data formats, such as CSV and SQL, but require preprocessing before use, including normalization of feature formats and cleaning of invalid data.

Currently, most carbon material databases focus on crystal structures and simulation-related data, such as crystal structures, adsorption energies, and electronic properties [50,51]. However, there is a significant lack of databases dedicated to the design processes of carbon materials (e.g., preparation methods, choice of activation agents, carbonization conditions) and their practical application performances (e.g., specific capacitance, adsorption capacity, catalytic efficiency). For instance, databases like NIST and CoRE include data on pore structures and adsorption isotherms, but these are predominantly centered on MOFs and zeolite materials. This gap limits the potential of data-driven research in advancing the design and performance optimization of porous carbon materials. In the future, it is imperative to develop databases specifically tailored to carbon material design processes and application performance. By integrating experimental data, simulation results, and process parameters, such databases could support the training of machine learning models and improve the efficiency of material design. Additionally, they would foster deeper research and broader applications of carbon materials in fields such as energy storage, catalysis, and adsorption.

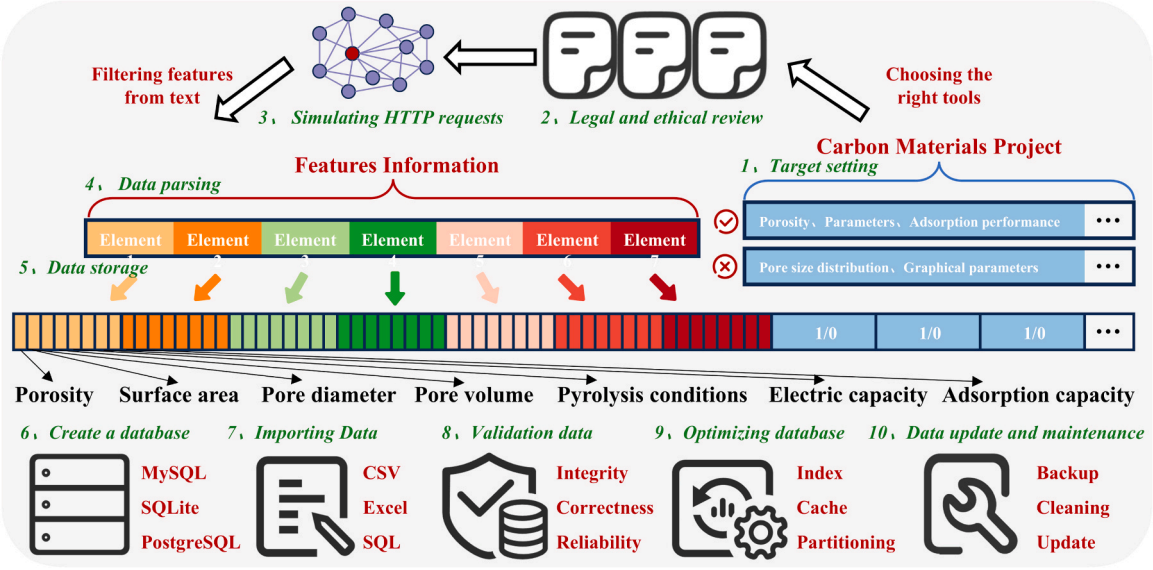


Fig. 4. The digital workflow for building a porous carbon materials database using web crawling techniques.

Table 2
Database information for related materials.

Database	Website	Material	Data Types	Quantity
PCOD	http://pcod.rsc.org/	Organic porous materials	Crystal structure, surface properties, specific surface area	22,000
Materials Projects	https://legacy.materialsproject.org/	Inorganic compounds	Crystal structure	-
OMDB-GAP1	https://omdb.mathub.io/	Large organic matter	Crystal structure, DFT band gap	12,500
MOFX-DB	https://mof.tech.northwestern.edu/	MOF	Structural data, adsorption properties	160,000
NIST-ISODB	https://github.com/NIST-ISODB/	Adsorbent materials	Adsorption performance	-
CoRE MOF	https://www.osti.gov/biblio/1118280	MOF	Structural data, adsorption properties	4700
CSD	https://www.psds.ac.uk/csd	Small-molecule organic	Crystal structure	1000,000
COD	https://www.crystallography.net/cod/index.php	Organic, inorganic, mineral	Crystal structure	513319
RCSR	http://rcsr.net/	Reticular crystals	Crystal structure	4115
DZS	https://www.iza-structure.org/databases/	Zeolite	Crystal structure, XRD	423
OQMD	https://www.oqmd.org	Inorganic materials	Crystal structure, energy, energy bands	471857
GDB−13	https://gdb.unibe.ch/	Organic small molecules	Molecular structure, topological structure, physicochemical properties	977 million
GDB−17	https://gdb.unibe.ch/	Organic small molecules	Molecular structure, topological structure, physicochemical properties	166.4 billion
ICSD	https://icsd.products.fiz-karlsruhe.de	Inorganic materials	Crystal structure, phase diagram information	307,301
IZA	https://www.iza-structure.org/	Inorganic materials	Crystal structure	100,000
NIST	https://www.nist.gov/	Chemical substances	Thermodynamic properties, spectral data, chemical reactions, crystal structures	Millions
MaterialGo	https://www.pkusam.com	Chemical substances	Crystal structure, physical, chemical properties	700,000
NOMAD	https://nomad-lab.eu/nomad-lab/	Chemical substances	Structural data, electronic properties, thermodynamic data, optical properties	3411628
Catalysis Hub	https://www.catalysis-hub.org/	Catalyst	Crystal structure, surface structure data, catalytic activation data	100000
CADS	https://cads.eng.hokudai.ac.jp/	Catalysts, Adsorbents	Crystal structure, pore data, adsorption properties, catalytic properties	-

3. Carbon material data processing

In recent years, ML has achieved significant success in establishing structure-property relationships for porous materials, effectively exploring vast materials spaces. Although ML models show great potential for predicting the performance of porous materials, their practical application is still limited by data availability [52]. A well-designed data processing workflow can significantly reduce noise in the data, enhancing its quality and usability, thereby better supporting the construction and application of ML models.

3.1. Data preprocessing

The quality of the source data has a significant impact on model performance, and improving the quality of training data is an effective way to enhance predictive capability [53].

3.1.1. Text data preprocessing

Selecting subsets of data that are highly similar to the specific prediction target, rather than using the entire dataset, can lead to better prediction results, as validated in the literature [54]. However, due to experimental conditions or human factors, data obtained from the literature may contain inaccuracies. The presence of outliers or erroneous data can interfere with the algorithm's pattern extraction, thereby affecting model accuracy, necessitating the removal of severely deviating data. Common outlier detection methods include Z-score, IQR, LOF, KNN, DBSCAN, isolation forest, and One-Class SVM [55,56]. Each of these techniques has distinct advantages: Z-score and IQR are effective for initial screening of simple numerical data; LOF and KNN are adept at identifying anomalies in local data distributions; DBSCAN works well for datasets with uneven density; and isolation forest and One-Class SVM excel in handling high-dimensional or non-linear data. These methods generally involve preprocessing the data, such as normalization, denoising, or filling missing values, to enhance detection accuracy. Anomaly scores are then calculated, or outliers are directly flagged, and potential anomalies are identified based on predetermined thresholds or scores. Depending on the application, the outliers can be removed, labeled as a separate category, replaced with mean or median values, or analyzed further to understand their causes, thereby ensuring a more reliable basis for subsequent data modeling and decision-making.

To ensure data reliability, it is essential to have a sufficient sample size while minimizing noise and bias. Additionally, the attributes used for modeling should have a reasonable dynamic range. If all data points have similar attribute values, the model will struggle to generate effective distinctions [57–59].

3.1.2. Graphics data preprocessing

Processing and analyzing image data of porous carbon materials is of significant importance in studying their structure-performance relationships. For experimental imaging data, preprocessing is a crucial step to ensure the reliability and consistency of data modeling and analysis. Denoising techniques, such as Gaussian filtering and wavelet denoising, effectively eliminate imaging artifacts, improving image quality. Resolution standardization ensures comparability across images, avoiding inconsistencies caused by scale differences. Feature extraction is an indispensable part of preprocessing, identifying key features like pore size distribution, shape irregularities, and surface area to lay the foundation for subsequent analysis. Advanced segmentation algorithms, including thresholding [60], edge detection [61], and machine learning-based methods, precisely extract regions of interest, such as pores or crystalline structures. For handling outliers, statistical methods or clustering-based algorithms, such as DBSCAN [62] and isolation forest [63], efficiently identify anomalous features, such as distorted or incomplete structures [64]. Data diversity and feature fidelity are critical for ensuring model validity. To prevent model bias and enhance its ability to capture structural features, sufficient sample

diversity must be ensured during preprocessing. In cases of limited data, data augmentation techniques can expand the sample size. A rational division of training, testing, and validation datasets provides a solid foundation for model development.

Different studies require tailored preprocessing strategies, but their common goal is to highlight the desired features while suppressing irrelevant or noisy features. For instance, Chan et al. [65] proposed a rapid microstructural characterization method combining topological classification with granularity distribution analysis. This method employs topological classifiers to identify local structures, voxelization for efficiency enhancement, and thresholding to improve feature distinction. Clustering and subdivision further refine microstructural feature extraction, enabling reverse mapping to atomic representations (Fig. 5b). Han et al. [66] applied preprocessing techniques such as labeling, color normalization, and data augmentation to optical microscopy (OM) images to identify material types and thicknesses in their study of 2D materials (Fig. 5c). For SEM images of porous carbon materials, pore shape feature extraction can involve steps like histogram equalization, edge enhancement, binarization, median filtering, and feature selection to ensure data clarity and consistency (Fig. 5a).

In the analysis of porous carbon material images, ML algorithms are increasingly applied. Classical models like ResNet [67], UNet [68], PixelNet [69], and AlexNet [70] excel in feature extraction and segmentation tasks. Models such as DefectSegNet [71] and DenseNet [72] are well-suited for microstructural feature extraction, while variational autoencoders (VAE) [73] and GANs [74] demonstrate outstanding performance in image generation tasks. Emerging approaches such as Swin Transformer [75] and Vision Transformer (ViT) [76] offer advantages in efficiently handling complex imaging tasks, particularly for pore feature extraction and microstructural recognition.

Supporting these algorithms are mainstream frameworks like TensorFlow, Keras, and PyTorch, alongside optimized platforms such as JAX, MindSpore, and PaddlePaddle [77]. These tools provide comprehensive development environments, pre-trained models, and support for large-scale distributed training, significantly improving the efficiency and quality of porous carbon material image analysis. In summary, scientific preprocessing methods, appropriate machine learning algorithms, and efficient framework selection are key to successful feature extraction and analysis of porous carbon material images. These methods enable researchers to delve deeper into the microstructural features of materials, uncover their structure-performance relationships, and provide theoretical and technical support for material design.

3.2. Feature engineering

After data preprocessing, it is essential to construct features (such as descriptors or fingerprints) that describe various characteristics of the materials, such as molecular structure, performance properties, or morphology. The construction of these features is closely related to the objectives of ML and relies on expertise and experience in the field of carbon materials [78]. The feature selection process includes feature extraction, feature analysis, and correlation and importance analysis. Features may exhibit correlations, and typically, the strength of these correlations determines whether to retain specific features [79]. For features with strong multicollinearity, generally only one is kept. However, when nonlinear correlations, significant interactions, or domain knowledge indicate their importance, multiple features are often retained to ensure data completeness [80]. For example, while BET surface area and pore volume are correlated, they cannot fully substitute for each other, so using micropore, mesopore, and macropore volumes

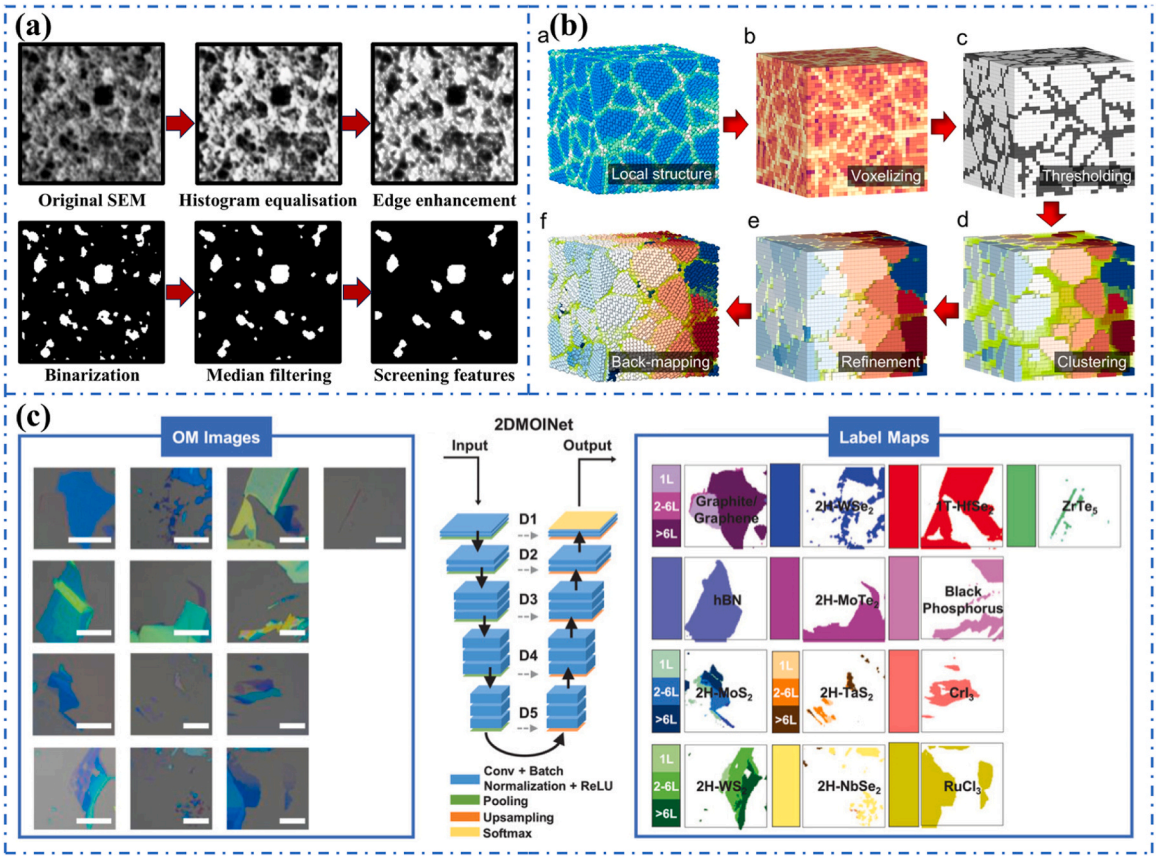


Fig. 5. (a) Schematic diagram of the SEM image preprocessing process of porous carbon materials. (b) Schematic diagram of the particle size distribution image preprocessing process. Reproduced with permission [65]. copyright 2020, Springer Nature. (c) Schematic diagram of the OM image preprocessing process. Reproduced with permission [66]. copyright 2020, John Wiley & Sons.

instead of total pore volume provides a more comprehensive representation of material characteristics. Additionally, an excessive number of features may lead to model bias or overfitting. For instance, the overlap between SSA and PV can introduce redundant information, affecting model performance.

To effectively select data features, three methods are typically employed:

1. Domain Knowledge-Based Feature Selection: Exclude features that do not affect the outcome based on experimental experience.

2. Correlation Analysis: Remove features with low correlation to the predicted outcome or those that are redundant, assessed using statistical methods such as Pearson correlation coefficient, Spearman correlation coefficient, mutual information, and chi-squared tests.

3. Automated Feature Extraction: Automatically extract complex features from raw data using hierarchical algorithms from deep learning, such as CNNs, autoencoders, DBNs, RNNs, GANs, and VAEs.

In practical modeling, the number of fitting variables should typically be less than half of the data points, and often even fewer. Therefore, downward selection of features and dimensionality reduction techniques are commonly used [81]. The effectiveness of feature selection may be influenced by hyperparameters, while dimensionality reduction reduces the number of features by projecting data from a high-dimensional space to a low-dimensional space. Sparse models generally have better generalization and interpretability, and can retain features that are highly relevant to the target [82]. During the dimensionality reduction process, removing redundant information and recombining features can effectively enhance the model's predictive capability, with model evaluation serving as a measure of the effectiveness of the reduction [83].

3.3. Data partitioning

The training process of ML involves iteratively training the model to bring its predictions closer to the true values. A proper division of available data is crucial for ensuring the reliability of the algorithm, typically splitting the data into training, testing, and validation sets. Table 3 lists some common data partitioning methods. In ML research related to carbon materials, cross-validation is one of the most commonly used methods, as it effectively enhances the model's generalization and robustness [84].

4. Models

ML, by utilizing training data and appropriate algorithms, deduces relationships within physical or chemical systems, serving as an empirical alternative to complex static and dynamic DFT and MD calculations. This approach provides deep insights into the relationships between nonlinear, multidimensional features [85]. It is particularly useful when the relationship between experimental variables and outcomes is unknown. In such cases, where the physical mechanisms are unclear or there is a highly complex dependence between variables and results, prediction methods based on physical laws may be difficult to implement or ineffective [86].

4.1. Algorithms

The selection of model algorithms must consider multiple factors, including objectives, data, and features, while integrating theories from fields such as quantum mechanics, phase-field theory, effective medium theory, and density functional theory. Table 4 summarizes algorithms commonly used for the synthesis and performance prediction of porous carbon materials. Different algorithms vary in how they define data relationships. In recent years, these algorithms have gradually been applied in the field of porous carbon material science. In Fig. 6, it is only in the past 5–6 years that ML has started to be used progressively in carbon material synthesis and application research. Over time, the accuracy of these models has steadily improved, accompanied by an increasing demand for larger datasets and greater computational power. Notably, 2016 marked the first instance of research involving robotic automation in controlling the design process of carbon materials [237]. In this study, an AI planner was connected to an automated growth reactor to feed back in situ characterization results to influence experimental procedures and achieve automated material design [130]. However, this area of research remains in its infancy, with significant room for development. Many models and methodologies are yet to be explored or optimized, presenting ample opportunities for future advancements in the automated and intelligent design of porous carbon materials.

4.1.1. Kernel model

Among kernel-based algorithms, Gaussian processes (GP) not only predict outcomes but also provide confidence intervals. However, their computational cost increases cubically with dataset size, making them suitable for small datasets with fewer than 1000 samples [87]. Another kernel-based algorithm is SVM, which has a significantly lower computational cost than Gaussian processes but does not provide confidence estimates for predictions [88]. These algorithms offer unique advantages in the design and synthesis of porous carbon materials. For example, Gaussian processes are particularly suitable for scenarios where data is scarce, as they deliver precise predictions while quantifying uncertainty through confidence intervals. This capability makes GP highly valuable for guiding experimental designs in material research, especially in cases where high costs or practical challenges limit the acquisition of large-scale experimental data [89]. On the other

Table 3
Data partitioning methods.

Method	Implementation process	Application Scenario	Advantage	Disadvantage	Extension Method
Cross-Validation	The dataset is divided into multiple subsets, one subset is the validation set, and the rest are the training sets, and this is repeated K times.	Small datasets, model selection	Less data waste, more stable model evaluation	High computational cost	Leave-one-out cross-validation
Hold-Out Method	Randomly split into the training set and test set.	Large datasets, fast model evaluation	Simple and direct	Model evaluation may be unstable and data is wasted.	Take the average of multiple
Bootstrapping	A new training set is generated by sampling with replacement, and the remaining data is used as the test set.	Small dataset, model stability evaluation	Providing evaluation for small datasets	The test set may not be representative enough, and there is a risk of overfitting.	Boosting
Leave-One-Out	Use one sample as the test set and the remaining samples as the training set, and repeat n times.	Accurate evaluation on small datasets	High accuracy	High computational cost	Leave-one-out cross-validation
Temporal Segmentation	The data is divided according to the time sequence to ensure that the time in the training set is earlier than that in the validation set.	Time-dependent data	Considering time dependencies	More complex and prone to information leakage	Expanding or rolling window method
Validation Set	Divide the data into training, validation, and test sets.	Used in model tuning	Avoiding overfitting	Validation set data waste	K-fold cross-validation

Table 4
Common algorithms for synthesizing porous carbon and predicting performance.

Classification	Algorithm	Advantage	Disadvantage	Applicable Scenarios
Small Data Methods	Linear Regression	Easy to understand	Sensitive to outliers and unable to handle nonlinear data	There is a linear relationship between the features and the target variable.
	Logistic Regression	Output probability, easy to interpret	Poor performance for nonlinear data and prone to underfitting	When the relationship between features is linear and interpretable
	Decision Tree	Easy to interpret, no need for much data preprocessing	Deeper trees are prone to overfitting	The data has an obvious hierarchical structure, or does not require feature scaling
	k-Nearest Neighbor	Simple and intuitive, no training phase required	Not suitable for large amounts of data and high-dimensional data, and has a high computational cost	Scenarios that require nonparametric models
	Gaussian Regression	Simple structure and few parameters	High computational complexity	When uncertainty estimates are needed, especially in time series and spatial data modeling
Big Data Methods	Neural Networks	Can learn complex nonlinear relationships	Requires a lot of computing resources and data, and the parameter adjustment is complex	The amount of data is huge and has high-dimensional characteristics
	Support Vector Machine	Performs well in high-dimensional spaces and is suitable for linear and nonlinear classification	Poor scalability for big data and multi-class classification problems	Handling high-dimensional datasets where the data is not linearly separable
	Random Forest	Strong resistance to overfitting, suitable for high-dimensional data, and easy to parallelize	The model is complex and difficult to interpret	When dealing with high-dimensional data and the risk of overfitting is high
	Naive Bayes	High computational efficiency, suitable for large-scale data, and easy to implement	Assuming that features are conditionally independent may lead to inaccuracy	Features are relatively independent
	Principal component analysis	Reduce data dimensions, help remove noise, and improve computing efficiency	The physical meaning of the features is not easy to interpret, and important information may be lost	Dimensionality reduction and noise removal of high-dimensional data

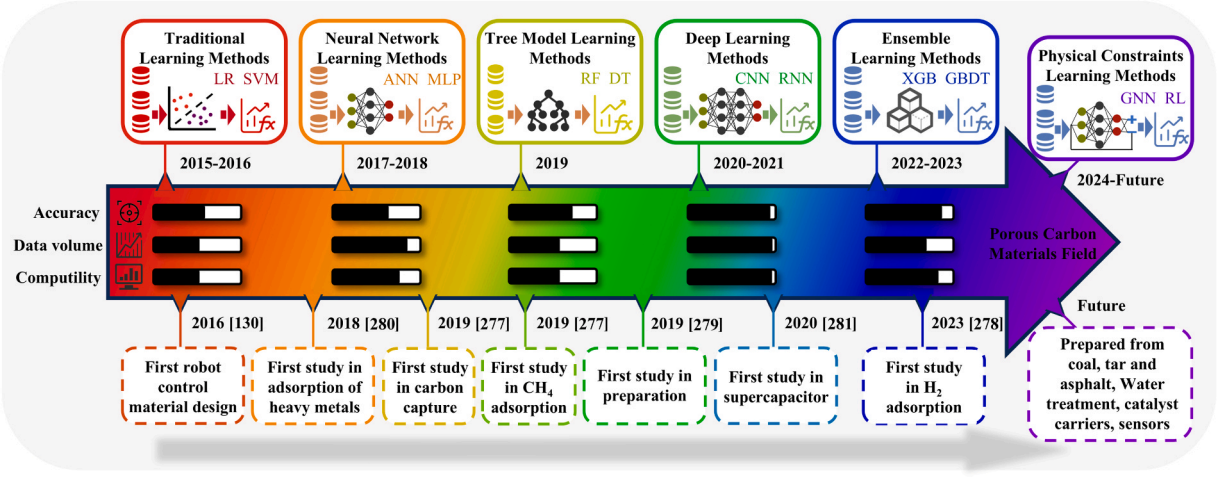


Fig. 6. Comparison of major milestones and algorithms in the application of ML in porous carbon materials [130,277–281].

hand, SVMs excel in rapid classification and regression tasks related to the synthesis conditions and performance metrics of porous carbon materials, owing to their lower computational requirements and robust generalization performance [90]. Leveraging the properties of kernel functions, both GP and SVM are adept at capturing complex nonlinear relationships between structure and performance in porous carbon materials. This ability provides deeper physical and chemical insights for material optimization. Consequently, these kernel-based methods hold significant value in addressing high-dimensional, small-sample problems in the study of carbon materials, offering efficient and interpretable tools for advancing material design and understanding.

4.1.2. Tree models

Tree-based algorithms (such as RF, GBDT, and DT) perform classification and regression tasks by constructing multiple decision trees and offer strong interpretability. DT recursively splits data features to form a tree structure, with nodes representing features and leaf nodes representing prediction outcomes. RF combines several independent decision trees to improve prediction performance and reduce overfitting through voting or averaging. GBDT is a gradient boosting algorithm based on decision trees, suitable for complex nonlinear problems, progressively reducing prediction error [91]. In the design and synthesis of porous carbon materials, these algorithms demonstrate unique advantages, especially when compared to kernel-based methods. One key strength of tree-based algorithms is their insensitivity to data distribution and feature scaling, eliminating the need for complex normalization processes. This makes them particularly well-suited for handling the heterogeneity and variability commonly observed in experimental data on porous carbon materials [92]. Moreover, tree-based models are good at dealing with mixed data types, such as the coexistence of numerical and categorical features, excelling in analyzing the multifaceted factors that influence material properties, such as synthesis temperature, raw material characteristics, and pore distribution. Another critical advantage is their interpretability. The hierarchical structure of decision trees

allows researchers to intuitively analyze the contribution of each feature to the performance of porous carbon materials. This transparency is invaluable for identifying key design parameters, optimizing synthesis processes, and exploring strategies to enhance material performance [93]. Additionally, the computational efficiency of RF and GBDT enables their application to larger datasets while maintaining high predictive accuracy in solving complex nonlinear problems. This capability is particularly advantageous in porous carbon materials research, where both experimental and simulation data often need to be integrated and analyzed. Through their robustness and interpretability, tree-based algorithms provide essential tools for advancing material design and achieving performance improvements.

4.1.3. Deep learning model

Deep learning algorithms rely on combinations of multiple nonlinear functions, and the multilayer stacking of neurons enables the model to learn data relationships ranging from simple features to complex patterns [94]. CNNs excel at processing image data, primarily extracting spatial features (such as edges, textures, and shapes of porous carbon) through convolutional layers, layer by layer, to obtain more complex abstract features, making them widely applicable in computer vision tasks [95]. CNN achieves feature extraction through sparse interactions, parameter sharing, and learning of local receptive fields. The architecture of a CNN is composed of alternating convolutional layers and subsampling layers. However, determining the specific CNN architecture for a particular application is a complex task, with various architectures reported in the literature to solve specific problems [96]. RNNs are suited for sequential data (such as text and time series), with their cyclic structure in the hidden layers allowing them to retain information from previous time steps, making them suitable for processing context-dependent data. Data from pyrolysis processes, adsorption curves, and the life cycle and capacitance variations of supercapacitors in carbon materials research align well with the characteristics of RNN, although their application in porous carbon materials is still relatively limited [97].

In the design and synthesis of porous carbon materials, deep learning algorithms exhibit unique advantages, creating a stark contrast with kernel-based and tree-based algorithms. Firstly, deep learning algorithms automatically extract complex, high-dimensional features from data without the need for manual feature engineering. This makes them particularly suitable for tasks such as analyzing fine textures, pore distribution patterns, and 3D reconstruction in porous carbon material imaging [98]. Compared to kernel-based algorithms and tree-based algorithms, deep learning models demonstrate superior performance in handling high-dimensional, nonlinear, and multimodal data [99]. Moreover, deep learning algorithms offer a unified framework for integrating diverse data types, such as images, time series, and textual information, in an end-to-end manner. For instance, CNN can predict the influence of pore size distribution on adsorption performance, while RNN can capture dynamic changes during pyrolysis processes. By combining these approaches, deep learning facilitates a holistic analysis of material performance. Additionally, deep learning's scalability makes it well-suited for leveraging large and growing materials databases and high-throughput experimental data. This capability allows for more efficient exploration of design spaces and the discovery of potential underlying patterns [100]. However, the high complexity of deep learning models poses challenges, such as significant computational resource requirements and reduced interpretability. Despite these limitations, their powerful learning capabilities provide unparalleled value in the intricate systems of porous carbon materials. Deep learning enables researchers to uncover nuanced structure-property relationships, optimize synthesis processes, and predict material performance, contributing to groundbreaking advancements in material design and development.

4.1.4. Reinforcement learning model

Reinforcement learning provides a general framework for creating artificial intelligence systems with reward and punishment feedback mechanisms, learning optimal strategies to achieve user-defined goals, such as experimental improvements [101]. Both Bayesian optimization and active learning can be viewed as special cases of reinforcement learning. Active learning aims to collect datasets that can train accurate ML models, while Bayesian optimization focuses on gathering data to identify optimal conditions for optimizing desired attributes [102]. The goal of Bayesian optimization is to optimize specific attributes (such as nanoparticle size), while active learning concentrates on building precise predictive models [103].

In the design and synthesis of porous carbon materials, RL demonstrates unique advantages and complements kernel-based, tree-based, and deep learning algorithms. By introducing dynamic decision-making processes and reward mechanisms, RL provides an innovative approach to material optimization. For instance, RL can guide experimental design by iteratively adjusting key parameters (e.g., pyrolysis temperature, precursor concentration, and activator ratio), learning optimal experimental conditions in real-time to enhance material performance (e.g., specific surface area or electrochemical stability) [104]. This stepwise optimization process is particularly suited for scenarios with high experimental costs or vast search spaces.

Additionally, Bayesian optimization has played a significant role in porous carbon material research. It can quickly identify optimal experimental conditions in complex design spaces. For example, in optimizing pore size distribution to improve adsorption performance, Bayesian optimization can quickly converge to an optimal solution with minimal experiments, significantly reducing the consumption of experimental resources [105]. In comparison, active learning focuses on building efficient predictive models. For instance, when studying the adsorption performance or conductivity of porous carbon materials, active learning selectively labels experimental data to progressively enhance model prediction capabilities. This method is especially effective when material databases are limited or experimental data acquisition is costly [106].

Overall, reinforcement learning and its derivative methods, with their dynamic adaptability and efficient utilization of experimental data, exhibit distinct advantages in the optimization and design of porous carbon materials. Particularly in scenarios requiring a balance between exploration and exploitation, RL's dynamic learning strategies stand in stark contrast to static algorithms such as kernel-based and tree-based methods. At the same time, RL addresses some limitations of deep learning when handling complex high-dimensional data. This multidimensional approach offers a robust solution for material optimization, enabling researchers to achieve breakthroughs in designing high-performance porous carbon materials.

4.2. Evaluation metrics

When building predictive ML models, a balance is typically needed between predictive capability, interpretability, and computational cost [107]. If the goal is to gain scientific insights, linear regression and tree-based models are more suitable due to their greater interpretability. Each model's computational cost varies based on implementation details and grows differently with dataset size, making dataset size a key factor in algorithm selection [108]. For instance, for a dataset of size A, the computational cost of a GP grows as A^3 , that of an SVM grows as A^2, that of a regression tree grows as Alog(A), and that of a neural network grows linearly with A. Therefore, deep learning models are particularly effective for large datasets, as they excel at capturing complex and nuanced relationships between experimental conditions and results [109]. In practical applications, it is common to test multiple algorithms and use the evaluation metrics shown in Table 5 to select the best-performing model. Selecting appropriate evaluation metrics is essential for assessing the performance of ML models accurately.

Table 5
Common evaluation metrics for synthesizing porous carbon and predicting performance.

Model Type	Evaluation Indicators	公式
Regression Model	Mean Squared Error	$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
	Root Mean Squared Error	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
	Mean Absolute Error	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$
	R-Square	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$
Classification Model	Accuracy	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
	Precision	$Precision = \frac{TP}{TP + FP}$
	Recall	$Recall = \frac{TP}{TP + FN}$
	F1 Score	$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$
	Confusion Matrix	$\begin{bmatrix} TN & FP \\ FN & TP \end{bmatrix}$
	ROC Curve AUC Curve	

4.3. Model interpretation

ML models can deliver accurate predictions, but how they reach these predictions is often unclear. This lack of interpretability makes extracting scientific insights from ML models challenging and may limit trust in these algorithms [110]. In predicting properties of porous carbon materials, black-box models are frequently used. To interpret these black-box models (such as DNNs, RFs, etc.), various methods are employed to reveal the models' inner mechanisms and help understand their decision-making processes.

In Fig. 7, common model interpretation methods were summarized and visually explained. The Shapley value is a significant approach based on game theory, which fairly distributes each feature's influence on the prediction. However, its high computational complexity makes it challenging to scale, particularly with high-dimensional data [111]. Local Interpretable Model-agnostic Explanations (LIME) perturb input features locally and use simple interpretable models (such as linear models) to approximate the behavior of complex models, providing localized explanations. In contrast, SHAP combines the advantages of Shapley values to offer a unified explanation framework, quantifying each feature's contribution to the prediction outcome. For example, Gradient-weighted Class Activation Mapping (Grad-CAM) visualizes the image areas that neural networks focus on, computing the regions of focus in the input image, revealing the basis for the model's decision-making. This is particularly useful for nonlinear neural networks [112]. Partial Dependence Plots (PDP) can show the average effect of a specific feature on the prediction result. However, they assume feature independence, which may introduce errors when features are correlated. Additionally, Counterfactual Explanations (CE) attempt to interpret model decisions by generating inputs similar to the current input but altered to change the prediction [113]. However, for complex models, this approach is computationally intensive, and generating logically consistent counterfactuals can be challenging [114]. These methods enhance our understanding of black-box model behavior, but high computational complexity and instability in local estimates remain major challenges when handling high-dimensional data or complex models.

5. Applications of ML models in porous carbon

In Fig. 8, machine learning optimizes the synthesis and application processes of carbon materials by learning patterns from data cases, providing precise guidance in selecting experimental conditions. It can analyze complex multivariable data, identify optimal parameter combinations, enhance material performance, reduce costs, and accelerate the development of new materials. Additionally, ML predicts synthesis

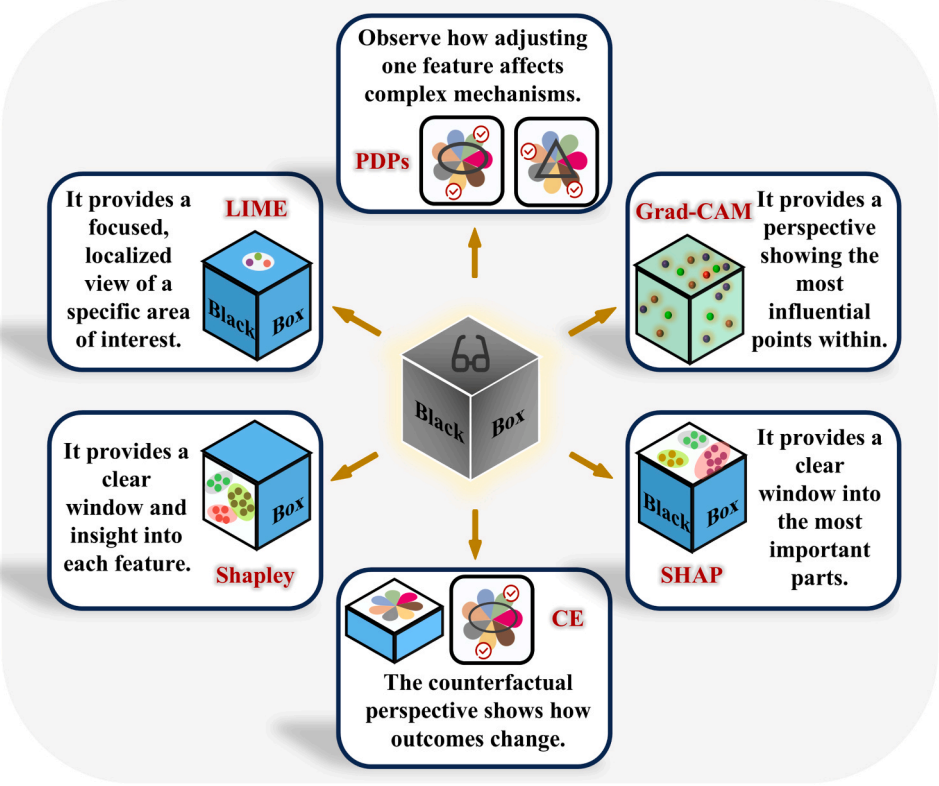


Fig. 7. Model interpretation methods and comparisons.

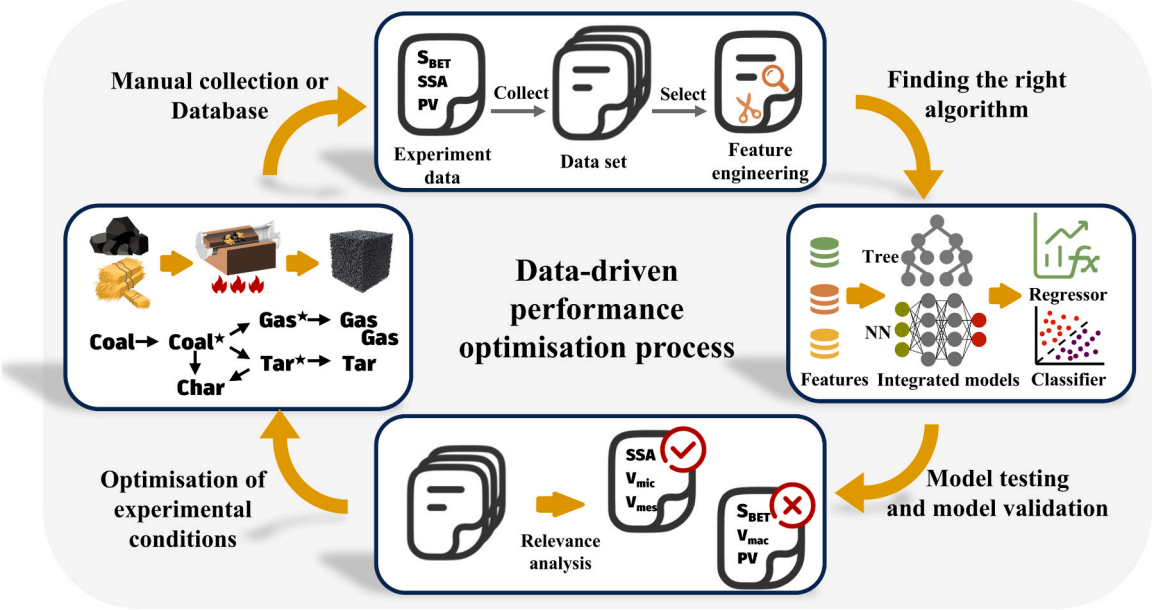


Fig. 8. ML framework for data-driven synthesis and application of porous carbon materials.

outcomes, minimizes the number of experiments needed, and fosters innovative applications of carbon materials in fields such as energy storage, catalysis, and adsorption (Table 6).

5.1. Design process of porous carbon based on ML

5.1.1. Design and optimization of experimental conditions

In the synthesis of porous carbon materials, various conditions (such as different temperatures) each exhibit unique characteristics. ML models can learn these features and ultimately integrate them across diverse experimental conditions to predict the structure or performance of porous carbon materials [115,116]. By summarizing patterns from past experimental data, ML models provide valuable guidance for the design and application of porous carbon materials. Theoretically, the SSA of defect-free, monolayer graphene carbon materials can reach 2630 m^2/g. By maximizing the exposure of sp^2-hybridized graphene ring surfaces and edges, this could theoretically increase to 7745 m^2/g [117]. However, as shown by the data collected in Table 2, the maximum specific surface area achieved in the laboratory so far is around 3500 m^2/g, which remains significantly lower than the theoretical value. This indicates that there is still considerable room for innovation and further exploration in this field.

The wide variety of raw materials for producing porous carbon, each with distinct compositional differences, adds complexity to predicting their formation processes. Although ML research on the formation of porous carbon is relatively limited, recent studies have started to focus on predicting the formation of porous carbon from biomass sources. Structural characteristics of porous carbon, such as porosity and SSA, have complex nonlinear relationships with pyrolysis conditions [118]. In addition, the noise and outliers in the prediction data due to experimental conditions and human factors further complicate the prediction model. However, the GBDT and GBDT show good robustness in this prediction task. These models use weak learners, with each learner focused on correcting the errors of the previous model, which effectively reduces overfitting. Using GBDT, inputs such as industrial and elemental analysis, biochemical composition, and pyrolysis conditions of biomass raw materials were used to predict the pore characteristics of porous carbon. Studies have shown that activation temperature and impregnation ratio are the main influencing factors, with nitrogen doping playing a significant role in micropore structure since NH_3 produced by the nitrogen source at high temperatures can enter micropores and expand pore sizes [119,120]. Similarly, the GBDT model also verified the criticality of the impregnation process and activation temperature to the performance of the final product [121]. As shown in Fig. 9, the feature importance rankings obtained from the GBDT and GBDT models indicate that the importance of features remains largely consistent between single and dual activation processes. Regarding experimental conditions, the impregnation ratio and activation temperature are the most influential factors affecting the SSA and PV of porous carbon materials. In terms of raw material characteristics, ash content and hydrogen content were identified as the most critical features. In another study, a tri-objective optimization GBDT model was used to predict the SSA and PV of biomass-based porous carbon materials. The study found that the impregnation process and pyrolysis conditions contributed approximately 44–57 % to the final structural performance, while the mixture of biomass components accounted for around 31–42 % of the influence. In contrast, the impact of elemental composition was less significant, contributing less than 20 %. This demonstrates that, in addition to the impregnation process and pyrolysis conditions, biomass composition mixing also significantly influences final performance [122,123]. It was found that Biomass mixed pyrolysis is more effective in producing products with excellent structural properties than single biomass pyrolysis. With their strong nonlinear handling capability and stability with unseen data, GBDT models have become ideal choices for predicting the pyrolysis process of porous carbon materials and are widely used in high-throughput calculations in fields such as force field charge distribution [124]. In the future, GBDT models can be combined with deep learning models to form integrated models. For example, GBDT could be used for feature engineering or an initial learning phase to generate features, which could then be input into deep learning models for higher-level abstraction modeling. Current studies on predicting the textural characteristics of porous carbon using material properties and pyrolysis conditions primarily focus on biomass raw materials. However, there is a wide range of raw materials for carbon materials, including coal, tar pitch, gasification slag, coke, fly ash, and residual carbon, which remain largely unexplored in ML prediction research.

The structural performance of porous carbon materials can be predicted not only based on raw material properties and pyrolysis conditions but also through features from the synthesis process or subsequent states. For example, electrochemical test procedures can quickly predict the SSA of carbon materials with an error margin of less than 14 %,

Table 6
Examples of ML applications in porous carbon-based materials.

Model	Data Source	Data volume	Scenario	Enter	Output	Hidden Layer	Neurons	Verification Method	Verification indicators	Refer.
MLP	literature	2794	CH_4 storage	SSA, V_{mic}, V_{mes}, temperature, pressure	CH_4 absorption、CO_2 and CH_4 selectivity	3	64、32、16	Cross validation	MSE 0.2、MAE 0.11、R^2 0.99	[182]
3D-CNN	Database	6500	CH_4 storage	Crystal structure data	CH_4 absorption	5		Leave-out method	MSE 0.003	[153]
RF(LR 、SVM、XGBT)	literature	1745	H_2 storage	Elemental analysis, V_{mic}, ultra-micropore volume, PV, SSA, pressure	H_2 adsorption			Cross validation	$R^2$0.9	[167]
DNN	Literature, experiment	1020	CO_2 Capture	Temperature, pressure, V_{mic}, V_{mes}, SSA	CO_2 absorption	2		Leave-out method、leave-one-out cross validation	Error 0.43	[149]
RF	literature	6244	CO_2 Capture	Elemental analysis, SSA, temperature, pressure	CO_2 absorption			Cross validation	R^2 0.919、MSE 0.187	[143]
DNN	Literature, experiment	1000	CO_2 Capture	V_{mic}, V_{mes}, SSA	CO_2 absorption	2	8、8	Cross validation、leave-one-out	MSE	[147]
DNN	literature	1452	CO_2 Capture	V_{mic}, V_{mes}, SSA, temperature, pressure	CO_2 and N_2 selectivity	2	7、7	Leave-out method、leave-one-out cross validation	R^2 0.96	[176]
CNN	Literature, experiment	970	CO_2 Capture	N2 adsorption isotherm at 77 K	CO_2 and N_2 selectivity	8	1024、256、1	Leave-out method		[177]
GBDT (LGB、XGB)	literature	632	CO_2 Capture	SSA, PV, V_{mic}, elemental analysis, adsorption parameters	CO_2 absorption			Cross validation	R^2 0.84、RMSE 0.66	[231]
Stack (XGBR、SVM、MLP)	Database		CO_2 Capture	Pore structure, catalyst correlation, DFT functional	CO、CO2和CO^{2-} adsorption energy	2	401	Cross validation	$R^2$0.97、MAE 0.19	[140]
CNN (GBR、MLP、LSTM、)	Database	527	CO_2 Capture	SSA, elemental analysis, PV, V_{mic}, temperature, pressure	CO_2 absorption	3	10、10、10	Leave-out method	R^2 0.857、MSE	[150]
RF	literature	1549	CO_2 Capture	Pore structure, chemical properties, adsorption conditions	CO_2 absorption			Cross validation	R^2 0.97、RMSE	[144]
MLR	literature	172	CO_2 Capture	S_{BET}, V_{mic}, temperature	CO_2 absorption			Cross validation	R^2 0.61、RMSE 1.05	[145]
XGB (ANN、SVM、MLP、DT、RF)	literature	2789	CO_2 Capture	Temperature, pressure, S_{BET}, pore size, PV, adsorption isotherm parameters	CO_2 absorption			Cross validation	R^2 0.998、MSE 0.0001、MAE0.003、RMSE 0.007	[142]
ANN	literature	288	Supercapacitors	S_{mic}, S_{mes}, N/O doping percentage	Specific capacitance	7	7 * 7	Leave-out method	MSE 38.5	[193]
RD(MLP、SVM、LR)	Literature, experiment	121	Supercapacitors	Potential window, SSA, PV, pore size, I_D/I_G, N-doping percentage, and O-doping percentage	Capacitance			Leave-one-out cross-validation	R^2 0.91、RMSE 0.68	[184]
GBR(DT、KNN、SVM、KRR、LR)	literature	2028	Supercapacitors	N-doping percentage, and O-doping percentage, I_D/I_G	SSA			Cross validation	R^2 0.95、MAE、RMSE	[131]
ANN (GLR、SVM、RF)	literature	70	Supercapacitors	Scan rate, S_{mic}, S_{mes}	Capacitor energy density, power density	1		Leave-one-out cross-validation	R^2 0.638、RMSE 0.957	[228]
ANFIS (LR、DT)	literature	260	Supercapacitors	ESR, RCT, PV, SSA, current density, potential window	Specific capacitance, power density			Leave-out method	R^2 0.9、MAE 39.76、RMSE 22.8	[192]
ANN(LR、XGB、RF)	literature	600	Supercapacitors	Potential window, SSA, pore size, ID/IG, N-doping percentage	Capacitance			Leave-out method	R^2 0.92、MAE 21.93、RMSE 29.32	[220]

(*continued on next page*)

Table 6 (*continued*)

Model	Data Source	Data volume	Scenario	Enter	Output	Hidden Layer	Neurons	Verification Method	Verification indicators	Refer.
ANN (SVMR、LR)	literature	260	Supercapacitors	Electrolyte materials, potential window, SSA, PV, Current density, ESR, RCT	Specific capacitance	1	60	Leave-out method	R 0.89、RMSE30.8、MAE46.6	[191]
XGB (MLR、ANN、SVN、GBM)	literature	105	Supercapacitors	S_{mic}/S_{mes}, S_{mic}/SSA, V_{mic}, V_{mes}, V_{mic}/V_{mes}, V_{mes}/PV	Capacitance			Leave-out method、cross validation	R^2 0.8、MAE 19.56、RMSE 25.5	[190]
MLP (CART)	literature	2997	Supercapacitors	Specific capacitance, cycle number, energy density, power density, PV	SSA, CA	3		Leave-out method	R^2 0.82、MAE 316.3、RMSE 420.7	[197]
LGB (DNN)	literature	195	Supercapacitors	Elemental analysis, Industrial analysis, structural composition, activation conditions, current density	Specific capacitance			Leave-out method	R^2 0.951、MAE 11.1、RMSE 14.76	[195]
ANN	Experiment	369	Design	Catalyst dosage, temperature, carbon source feed	Yield, quality, structure, photoelectric quality factor	3		Cross validation	MAPE 27 %、23 %、22 %、3.6 %	[206]
GBDT(RF、XGB)	literature	258	Design	Industrial analysis, elemental analysis, biochemical composition, pyrolysis conditions, nitrogen content	SSA, PV, microporosity			Cross validation	R^2 0.939、RMSE 216.9	[232]
GBR(RF)	literature	169	Design	Biomass characteristics (composition by industrial analysis), pyrolysis conditions (temperature, time, heating rate)	SSA, PV, yield			Cross validation	R^2 0.915、RMSE 4.25	[235]
GBDT(RF)	literature	131	Design	Elemental analysis, industrial analysis, impregnation conditions, pyrolysis conditions, activation conditions	SSA, PV, yield			Cross validation	R^2 0.91	[234]
EI-SVM	Experiment	100	Design	Catalyst composition, reaction temperature, time, atmosphere	Carbon nanotube diameter, yield				MAE 0.01	[137]
CNN	Experiment	1000	Defect Analysis	Nanoparticle characteristics obtained by SEM of carbon material with attached nanoparticles	Nanoparticle coordinates, order			Cross validation		[254]
Ensemble Method	Simulation	3000	Pollutant Removal	SSA, ultra-micropore + S_{mes}	Paracetamol adsorption			Cross validation	$R^2$0.97、RMSLE	[230]
SVM	Experiment	399	Gas separation	FFV, average interlayer spacing of graphite-like carbon sheets, carbonization temperature	Breathability, separation			Cross validation	R^2 0.794、MAE 0.139、RMSE 0.281	[238]
KRR	literature	2000	Binding Energy	CHO material structure data, DFT data, GW data	XPS spectra			Cross validation	MAE、RMSE 0.608	[194]

effectively avoiding the need for complex testing procedures [125]. In additive manufacturing, ML can detect porosity sizes in melt pool images with an error rate of less than 0.01 %, replacing high-cost and time-intensive porosity testing [126]. In the future, well-trained ML models may uncover the necessary conditions for synthesizing novel porous carbon materials or optimize synthesis conditions to achieve higher porosity, larger micropore volumes, or increased yields. The development of such models stands to gain significantly from both successful and failed experiments [127]. However, due to the limited availability of reported failure data in the literature, encouraging researchers to publish these results could support more comprehensive model training and advancement.

5.1.2. *Design and optimization of molecular structure*

In the molecular structure design of porous carbon materials, the combination of traditional computational simulations and ML methods has demonstrated tremendous potential. Researchers have successfully constructed structured models of carbide-derived carbons using techniques such as annealing MD simulations (Fig. 10a) [128], all-atom MD simulations (Fig. 10b) [129], MC simulations (Fig. 10c) [130], particle

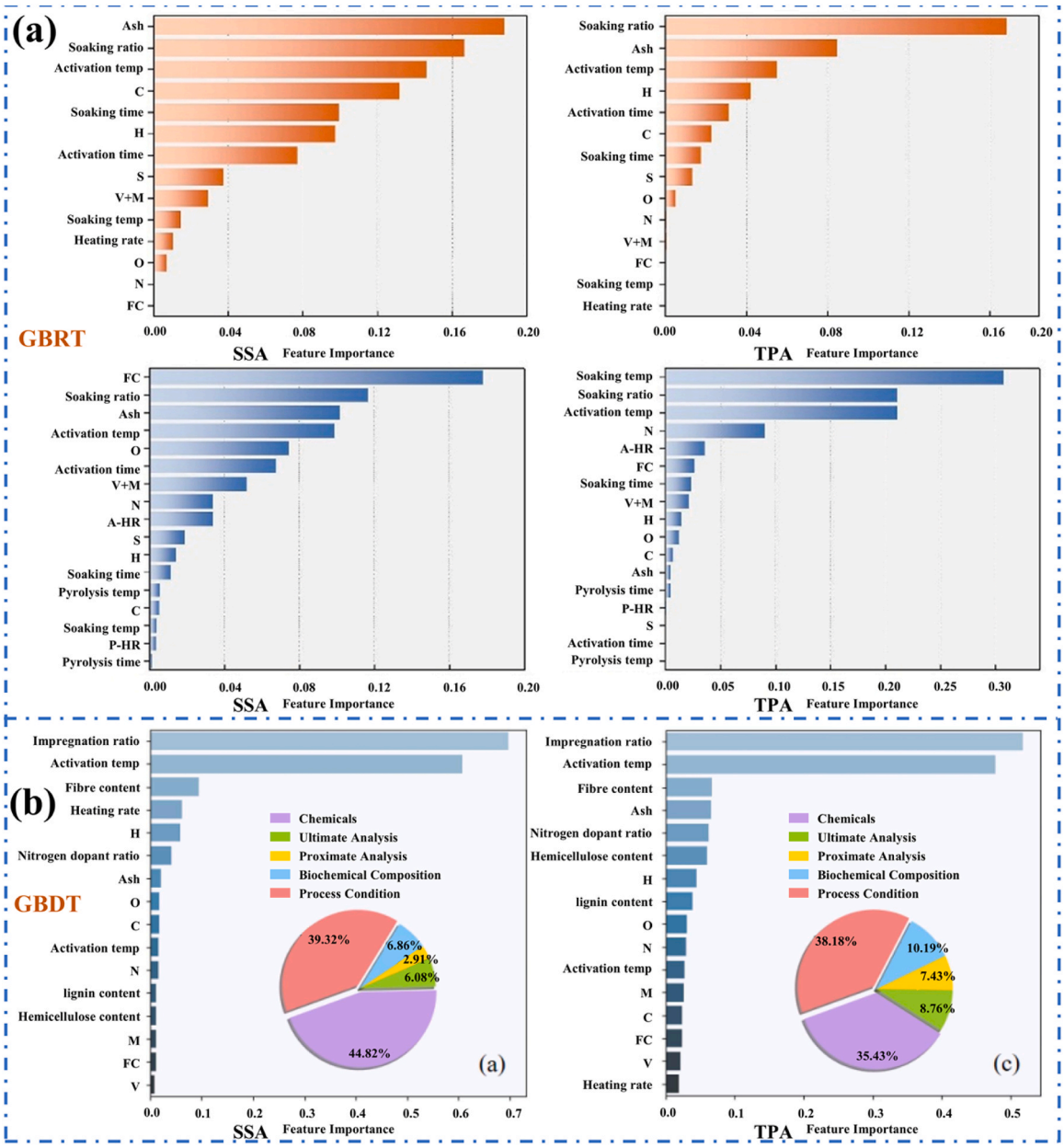


Fig. 9. (a) Feature importance rankings of the GBDT model for single-step and two-step activation of biomass-based porous carbon. Reproduced with permission [121]. copyright 2024, Elsevier. (b) Feature importance rankings of the GBDT model for nitrogen-doped biomass-based porous carbon. Reproduced with permission [119]. copyright 2024, Elsevier.

swarm optimization, and cooperative assembly methods. However, these approaches are often computationally expensive, especially when dealing with large-scale simulation systems. As a result, ML methods have recently been introduced to assist in optimizing traditional simulations and material design.

In this context, Wang et al. [131,132] integrated ML-driven Gaussian Approximation Potential (GAP) models with MD simulations to conduct in-depth studies on the microstructure and pore size distribution of porous carbon materials. ML techniques were employed to develop precise interatomic potentials, accelerate large-scale simulations, quantify pore size distributions and mechanical properties, and optimize material design. For instance, the GAP model enabled accurate prediction of vacancy defect formation energies in graphene sheets, facilitating large simulation units with 131,072 atoms using GAP force fields (Fig. 10e). Additionally, ML was applied to analyze the ratio of sp^2 to sp^3 bonds in graphene sheets, providing insights into the distribution patterns of different carbon rings within the topological structure. The study revealed that as material density decreases, pore size increases and exhibits a broader distribution, but mechanical properties decline. ML methods significantly reduced computational costs, offering an effective tool for designing and optimizing porous carbon materials with tailored pore characteristics.

Based on the GAP model and pore size distribution predicted by ML, the theoretical storage capacity of porous carbon materials for hydrogen, carbon dioxide, methane, acetone, etc. [132] can be calculated. Leveraging ML techniques significantly improved the accuracy of predictions while reducing the computational resources required by traditional simulations. Zhang et al. [133] further expanded the application of ML by integrating MD simulations and GBR algorithms to construct interface models of porous carbon electrodes/electrolytes with varying pore size gradients (Fig. 10d). By predicting the specific capacitance under different pore size combinations, they found that the

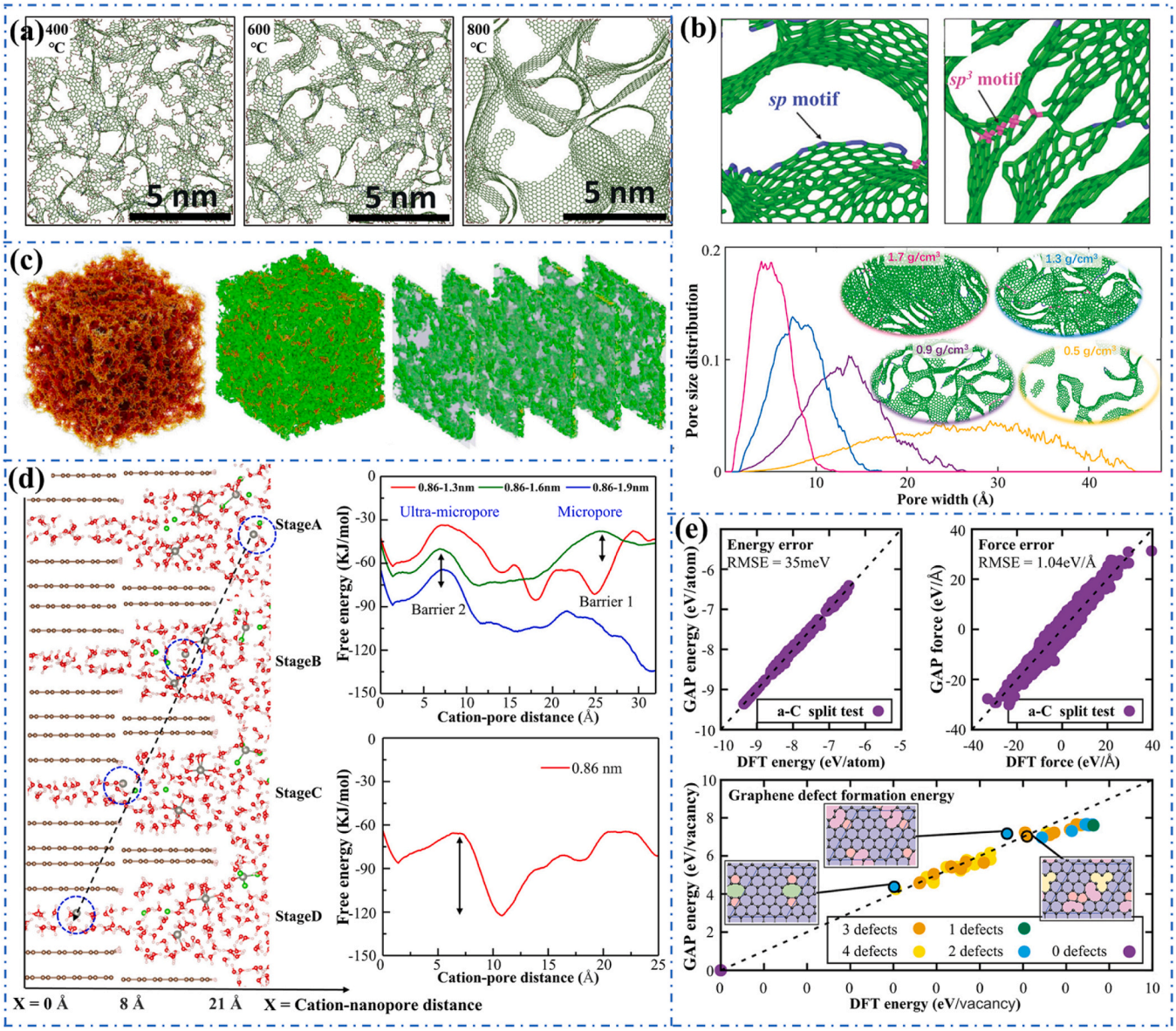


Fig. 10. (a) Porous carbon model constructed by annealing the MD simulation. Reproduced with permission [128]. copyright 2024, Royal Society of Chemistry. (b) Porous carbon model constructed by all-atom MD simulation. Reproduced with permission [129]. copyright 2022, American Chemical Society. (c) Porous carbon model constructed by Monte Carlo simulation. Reproduced with permission [130]. copyright 2024, American Chemical Society. (d) Porous carbon electrode/electrolyte interface model with different pore size gradients and the storage process of Zn^{2+}. Reproduced with permission [135]. copyright 2025, American Chemical Society. (e) Accuracy of GAP prediction for graphene with different vacancy defects. Reproduced with permission [129]. copyright 2022, American Chemical Society.

combination of ultra-micropores (0.6–0.9 nm) and micropores (1.6 nm) achieved the highest specific capacitance. This result provides crucial theoretical guidance for the optimized design of porous carbon in energy storage applications.

The application of ML in the molecular structure design and optimization of porous carbon materials demonstrates tremendous potential. From developing efficient interatomic potentials to accelerating large-scale simulations and data analysis, and integrating traditional methods to optimize pore characteristics, ML methods exhibit distinct advantages. Future research directions include integrating ML with quantum mechanics simulations, molecular dynamics, and continuum models to achieve multiscale material design. Additionally, using large datasets generated by high-throughput computations to train models that predict porous carbon materials with novel pore characteristics, as well as introducing reinforcement learning methods for adaptive design of complex porous structures, are promising areas to explore. In conclusion, ML opens new pathways for the molecular structure design and performance optimization of porous carbon materials. It not only significantly reduces computational costs but also enhances design efficiency and accuracy, offering significant value for the future development of porous carbon materials.

5.1.3. *Design and optimization of pore structure*

In the design of porous carbon materials, ML has emerged as a pivotal tool for overcoming the limitations of traditional methods. Precise control of pore structures requires the synergistic optimization of multiple parameters, including pore size distribution, pore connectivity, and surface chemistry, which often exhibit complex nonlinear relationships. Studies have demonstrated that integrating high-throughput characterization data with ML algorithms enables the inverse design of hierarchical pore structures. For instance, GBDT models can effectively analyze the correlation between process parameters (e.g.,

activation temperature and impregnation ratio) and pore size distribution [121], while CNN and DNN are better suited for processing unstructured data, such as SEM images, to achieve intelligent pore morphology recognition and classification [282]. In the context of mesopore-macropore co-design, the combination of MD simulations and ML-based interatomic potentials significantly improves the accuracy of pore connectivity predictions, even enabling the construction of carbon materials with tailored pore size combinations (e.g., ultra-micropores of 0.6–0.9 nm and micropores of 1.6 nm) [133]. ML models can capture cross-scale structural features that are challenging for conventional simulation methods to resolve.

The precise regulation of surface functional groups represents another critical dimension in optimizing pore structure performance. By analyzing extensive XPS data, ML approaches can predict the distribution patterns of oxygen- and nitrogen-containing functional groups on carbon materials and reveal strategies for modulating their surface-active sites [283]. Recent studies have also shown that reinforcement learning can dynamically optimize chemical vapor deposition (CVD) process parameters, enabling fine control over surface chemistry. The method is particularly valuable for fabricating functionalized porous carbons with tailored catalytic properties [284]. Notably, the mechanical stability of pore structures must not be overlooked. GNN-based predictive models can assess the deformation behavior of different topological structures under stress, offering new insights for developing porous carbons with both high SSA and superior mechanical performance [285].

For complex three-dimensional pore networks, extracting features such as pore geometry, structural robustness, and grading hole remains a task uniquely suited to ML, yet one that still presents challenges. Additionally, establishing interpretable mappings between pyrolysis conditions and pore structure performance requires the development of novel algorithmic frameworks. Future research will likely focus on hybrid ML models that incorporate physical constraints, such as thermodynamic principles and transport phenomena, to enhance generalization across unexplored material systems. Further advancements may also leverage digital twin technology to enable multiscale optimization, from atomic-level design to macroscopic performance prediction.

5.2. Gas adsorption

5.2.1. Carbon capture

In carbon capture research, ML is widely used to predict the adsorption performance of porous carbon materials, demonstrating remarkable adaptability [134]. The basic mechanism of carbon capture relies on van der Waals forces between the carbon material and CO_2 molecules, which restrict the movement of CO_2 molecules [135,136]. The level of adsorption energy is crucial to capture efficiency: higher adsorption energy increases CO_2 capture capacity, but if it is too high, it may hinder desorption. Therefore, accurately predicting adsorption energy is essential [137–139]. The Stacking model that combines the MLP, SVM, and XGBR algorithms can predict the adsorption energy of CO and CO_2 in different systems. In particular, the calculation based on the RPBE functional theory works best in this model [140,141]. This method enables fast predictions of adsorption energy for ionic species on metal surfaces—results that cannot be directly obtained from DFT calculations.

Using the texture characteristics of carbon materials and adsorption conditions, ML can precisely predict carbon capture performance, significantly reducing experimental costs and providing an efficient pathway for new material screening and process optimization, which promotes the industrialization of carbon capture technologies. An RF model has been established to analyze the impact of pressure, pore distribution, and textural characteristics on CO_2 adsorption capacity. Studies indicated that the material's SSA, mesopore volume, and micropore volume contribute most significantly to CO_2 adsorption capacity, with ultra-micropore volume being especially impactful [143]. It was also observed that under low pressure (0–0.2 bar), pressure and pore structure influence 80 %-90 % of the carbon capture capacity. Meanwhile, at higher pressures (0.6–1 bar), chemical composition and pore structure account for 72 %-95 % of the influence (Fig. 11a_2 and a_4). Among chemical composition features, nitrogen groups are the most influential on carbon capture, followed by oxygen and carbon content. At low pressures (0–0.15 bar), the impact of nitrogen groups is especially pronounced (Fig. 11a_2). Regarding pore structure characteristics, micropore volume has the greatest influence on carbon capture [144], followed by SSA. At higher pressures (0.15–1 bar), ultra-micropores play a critical role (Fig. 11a_3). As pressure increases, the contribution of micropore volume to carbon capture performance slightly decreases, while the influence of mesopore volume significantly increases. This indicates that the range of pore sizes capable of adsorbing CO_2 expands with increasing pressure (Fig. 11b).

Additionally, feature extraction using DT, XGB, and ANN models suggests that materials with a combination of macropore volume (0.5–1.45 cc/g), high specific surface area (>450 m^2/g), and ample ultra-micropores (<0.442 nm) show optimal CO_2 adsorption performance [145]. In adsorption performance prediction, the textural characteristics of micropores, mesopores, and macropores are all essential, with micropore-dominated pore distributions performing best. For operational conditions, studies indicated that optimal carbon capture performance is achieved at low temperatures (around 20 ℃) and pressures above 1 bar [146].

Deep learning models have shown significant advantages in predicting carbon capture capabilities, enhancing both accuracy and speed of predictions while adeptly handling complex nonlinear relationships and automatically extracting key features [142]. By using inputs like micropore volume, mesopore volume, and S_{BET}, DNN models have been applied to predict CO_2 adsorption capacity in porous carbon materials, aiming to assess CO_2 capture potential based solely on textural features [147,148]. Research indicated that micropores are crucial for CO_2 adsorption, while a small number of mesopores are more effective in aiding CO_2 diffusion through the pore channels. In such predictions, S_{BET}, as a self-regulating structural parameter, is particularly effective in representing gas-solid interactions, thereby improving adsorption capacity predictions when combined with other structural parameters [149]. Comparing various types of models to predict carbon capture effects, it was found that tree models (RF and XGBoost) and deep learning models had the highest accuracy (Fig. 11c). Comparative studies of deep learning algorithms (CNN, MLP, LSTM) reveal that CNN is the most sensitive to carbon capture features, delivering the highest prediction accuracy [150,151]. Temperature and pressure emerge as primary influencing factors, followed by textural features, carbon-to-temperature ratio, and carbon-to-pressure ratio. Using GBDT to predict the CO_2 adsorption performance of biomass-derived porous carbons showed that the characteristics that affect the adsorption performance are adsorption parameters, structural characteristics, and composition characteristics, in order. In particular, the pore volume in the pore size range of 0.6–0.8 nm plays a key role in the efficient CO_2 capture process. This also shows that for porous carbon materials that are not doped with nitrogen and oxygen, the raw material composition characteristics are only a secondary factor [152].

With the continuous advancement of ML and deep learning technologies, the field of carbon capture is poised to embrace more intelligent and efficient solutions. By integrating multidimensional data (such as pore structure, pore shape, chemical composition, operating conditions, and kinetic behavior) into comprehensive predictive models, researchers can not only optimize material design but also provide reliable guidance for process parameter adjustments. In addition, ML-based inverse design is expected to become a vital tool for the development of carbon capture materials. By inputting target performance metrics, models can rapidly screen or design material combinations that meet specific requirements, significantly shortening the development cycle for new materials. This approach holds immense potential for exploring

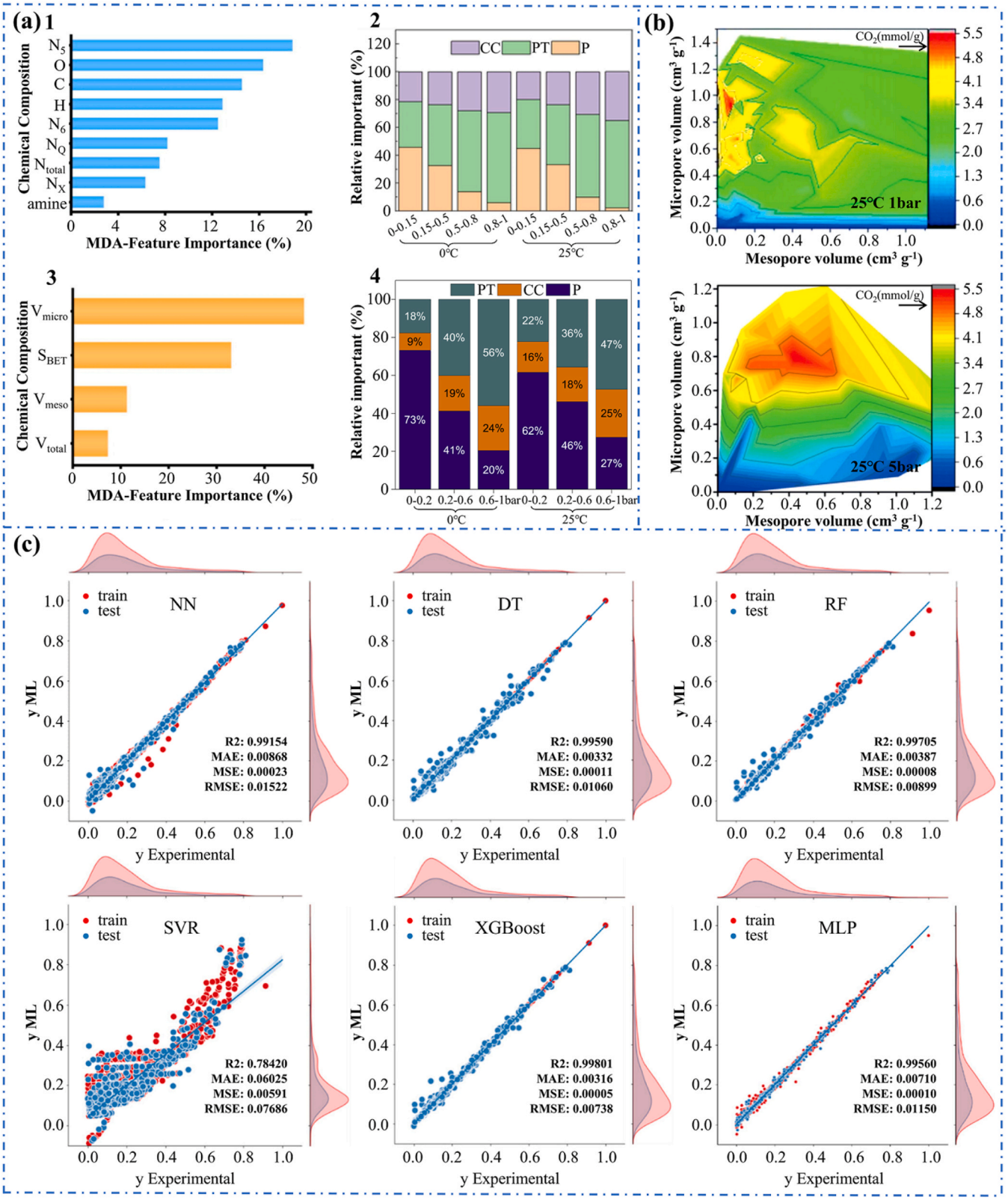


Fig. 11. (a) Importance of features for predicting carbon capture performance. a_1 and a_3, Reproduced with permission [146]. copyright 2024, Elsevier. a_2, Reproduced with permission [144]. copyright 2023, Elsevier. a_4, Reproduced with permission [143]. copyright 2020, Elsevier. (b) Correlation between pore structure and carbon capture performance at different pressures. Reproduced with permission [149]. copyright 2022, Springer Nature. (c) Accuracy of several models for predicting carbon capture performance. Reproduced with permission [142]. copyright 2024, American Chemical Society.

novel porous structures, functional modifications, and composite material systems. Moreover, combining reinforcement learning with process optimization for carbon capture can dynamically adjust experimental conditions to maximize capture efficiency. Future research could further integrate automated experimental platforms with intelligent algorithms to achieve end-to-end automation (from material design and experimental validation to performance optimization), thereby accelerating the industrial application of carbon capture technologies.

5.2.2. *Methane storage*

Methane storage in porous materials involves both the microstructure and macroscopic performance of the materials, typically through physical adsorption [153]. MLP models, with their deep architectures, can capture multi-scale information and link microscopic and macroscopic features through combinations of nonlinear activation functions and layered networks, producing a comprehensive description of storage behavior. Using backpropagation and gradient descent optimization, MLP effectively handles noise and captures the associations among nonlinear features, enhancing robustness to experimental data. Additionally, neural network models achieve greater accuracy and computational efficiency compared to Monte Carlo (MC) predictions. For instance, calculating adsorption isotherms via MC takes over 56 CPU hours, whereas CNN computes them in under 0.02 seconds (excluding grid generation) [154]. In methane absorption prediction, MLP models used SSA, micropore volume, mesopore volume, temperature, and pressure as inputs, showing high accuracy [155]. Due to the similarity in properties between CO_2 and CH_4, the key features influencing methane storage align closely with those of carbon capture and will not be reiterated here. However, because CO_2 and CH_4 differ in molecular dynamic diameters, the pore sizes critical to methane storage differ from those for carbon capture (Fig. 12a). Compared to carbon capture, mesopore volume plays a more significant role in methane storage at any pressure, as the molecular dynamic diameter of CH_4 is approximately 1.2 times larger than that of CO_2 [156]. Additionally, as pressure increases, the contribution of micropore and mesopore volumes also increases, although this trend plateaus when the pressure exceeds 10 bar (Fig. 12c). When comparing eight models (KNN, LASSO, Elastic Net, SVM, XGBoost, RF, MLP, and ANN) for methane storage performance prediction, the latter four models exhibited the highest accuracy (Fig. 12b). Moreover, as dataset size increases, the prediction accuracy of ANN and MLP models further improves [157,158]. Notably, removing SSA from input features weakens the nonlinear relationship, reducing pore volume to a simple direct proportionality with the prediction [159]. Adding SSA reintroduces complex nonlinearity, indicating that

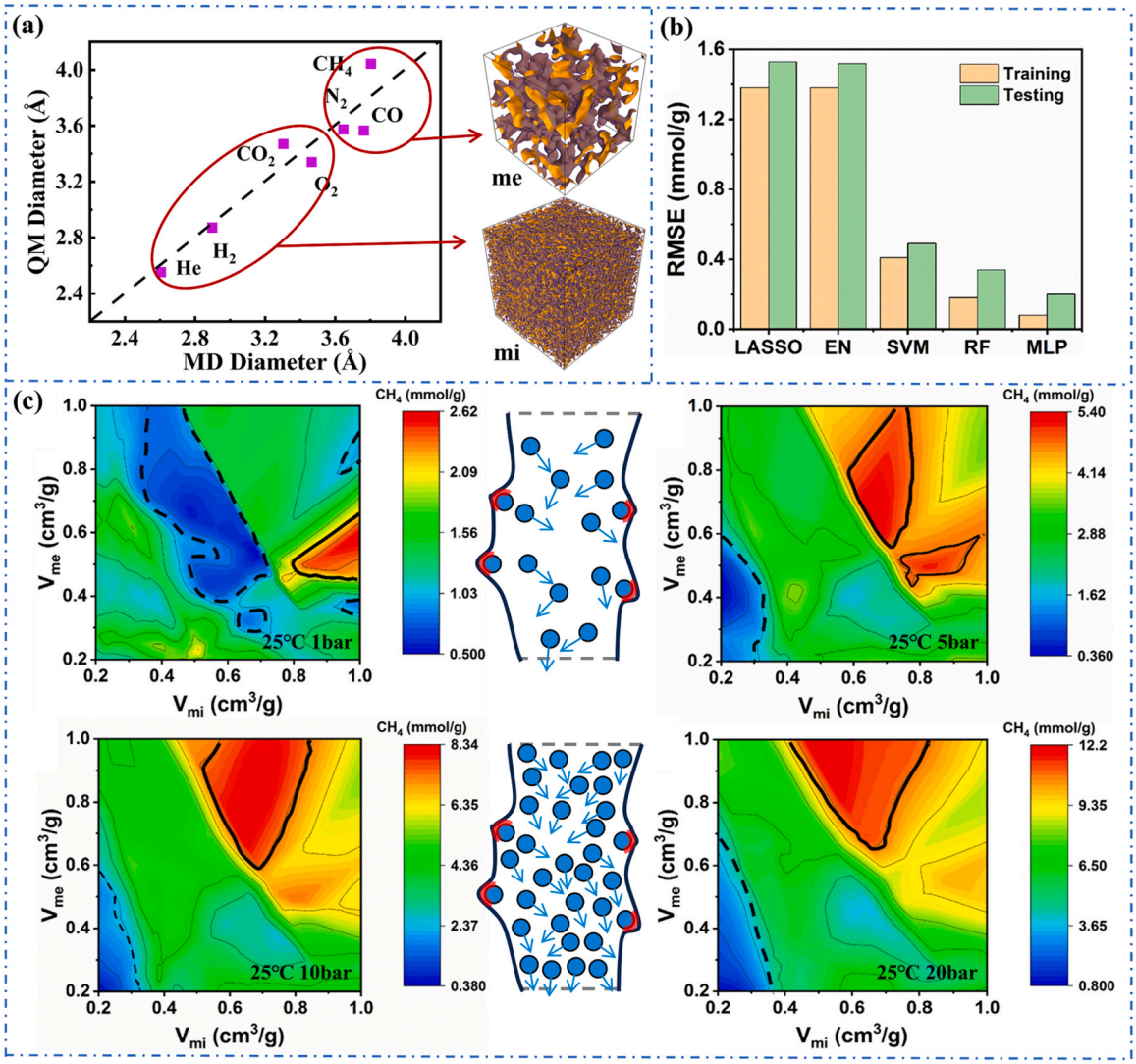


Fig. 12. (a) Quantum mechanical diameter and molecular dynamics diameter of gas molecules. (b) Accuracy of the model for predicting methane storage performance. Reproduced with permission [155]. copyright 2021, Elsevier. (c) Correlation between pore structure and methane storage performance at different pressures. Reproduced with permission [155]. copyright 2021, Elsevier.

while SSA correlates with micropore volume and mesopore volume, it likely encompasses additional complex features, which the authors suggest may be linked to pore irregularity. Experimental data strongly support this finding: The AC-Mix material with micro-mesoporous heterostructure (exhibiting a SSA of 869.82 m^2/g) demonstrates significantly superior methane adsorption capacity compared to single-pore-structure activated carbons [160]. This phenomenon not only confirms the correlation between SSA and PV but also suggests that SSA may encompass complex structural characteristics such as pore irregularity, which is in perfect agreement with the nonlinear contribution of SSA observed in machine learning models.

In the field of methane storage using porous materials, the application of deep learning models can significantly advance research and development. First, advanced algorithms such as GAN and VAE can be employed to generate virtual samples with specific pore structures. These virtual datasets can compensate for the scarcity of experimental data and optimize the material design process [161]. Second, the development of Physics-Informed Neural Networks (PINN) can embed thermodynamic and kinetic mechanisms of gas adsorption into the model. This enhances the physical credibility and interpretability of the predictions, bridging the gap between data-driven methods and domain knowledge.

5.2.3. Hydrogen storage

In the hydrogen storage process of porous carbon materials, both physisorption and chemisorption may occur [162]. Chemisorption, driven by strong binding enthalpy, leads to the dissociation of hydrogen molecules into atoms, which then bind to the surface of the carbon material. In contrast, physisorption primarily relies on van der Waals forces, adsorbing hydrogen molecules or atoms onto the material's surface [163]. The loading of heavy metals on carbon materials can catalyze the dissociation of hydrogen molecules, making the resulting hydrogen atoms more readily adsorbed [164]. Additionally, metal loading alters the energy state of the carbon material's surface, improving hydrogen storage efficiency at low temperatures [165]. Researchers have used a Shapley Additive Explanations (SHAP) algorithm, based on game theory, to analyze the relative importance of variables in hydrogen storage prediction. Recently, this method has gained popularity in various ML models for its strong interpretability [166]. Kusdhany et al. [167,168] analyzed SHAP values for features influencing excess hydrogen adsorption and identified pressure as the most significant experimental factor. They also observed a positive correlation between excess hydrogen adsorption and oxygen content. Specifically, within the range of 8–11.5 wt% oxygen content, excess hydrogen adsorption increased by 0.6 wt%. However, adding more oxygen beyond 11.5 wt% showed diminishing benefits (Fig. 13a_1). Furthermore, at higher oxygen contents (indicated by purple and red points), the hydrogen content has a smaller positive interaction with excess hydrogen uptake (Fig. 13a_2). This finding aligns with the experimental results of Gotzias and Schaefer [169,170], who reported that hydroxyl and carboxyl groups are more strongly correlated with excess hydrogen adsorption compared to other epoxy functional groups. Regarding pore structure characteristics, total pore volume and micropore volume showed a linear positive relationship with excess hydrogen adsorption (Fig. 13a_3 and a_7), while ultra-micropore volume exhibited a saturation curve (Fig. 13a_5). Specifically, ultra-micropores (<1 nm) demonstrated greater impact on excess hydrogen adsorption at 77 K before reaching a saturation point, beyond which additional ultra-micropore volume provided negligible benefits. Both experimental and theoretical studies indicate that at 77 K, pore widths between 0.5 and 0.7 nm are optimal for hydrogen adsorption due to the overlapping potential fields of the pore walls [171]. At lower pressures, the overall influence of all pore volume types on hydrogen adsorption is weak. However, at higher pressures, micropore volume maintains a significant linear contribution to excess hydrogen adsorption (Fig. 13a_8), while the effects of total pore volumes and ultra-micropore volumes diminish (Fig. 13a_4 and a_6). For features derived from DFT simulations, binding energy, metal charge transfer, and second ionization energy were identified as the top three most important characteristics, collectively accounting for 64.6 % of the total feature importance (Fig. 13c). Among various ML models compared, RF and GBR achieved the highest prediction accuracy for excess hydrogen adsorption in porous carbon materials (Fig. 13b).

The weak hydrogen bonds in carbon materials result in lower adsorption energy, which is not conducive to hydrogen adsorption within narrow pore channels. Therefore, increasing the available surface area is closely related to enhancing hydrogen storage capacity [172, 173]. Existing studies suggest that the adsorption performance of carbon materials is primarily linked to micropore or ultra-micropore volume. Additionally, an appropriate amount of mesopores and macropores is also essential, as they improve gas flow within the carbon material. Small pores provide more wall contact, enabling stronger chemical bonds or intermolecular forces between the adsorbed molecules, thus enhancing adsorption efficiency [174]. Doping porous carbon materials with elements like oxygen, nitrogen, or metals can increase surface active sites, enhancing chemical bonds and intermolecular interactions, thereby improving adsorption capacity. Furthermore, irregular pore shapes and protruding structures may also impact adsorption capacity. Pores with more protrusions and greater irregularity could provide a larger adsorption surface area, further enhancing adsorption capacity. The effect of this structural characteristic on adsorption performance has not been extensively studied and warrants further exploration. Future research on porous carbon materials for hydrogen storage should focus on multiscale synergistic optimization and mechanistic insights. By integrating experimental methods with advanced machine learning models (such as RF, GBR, and deep learning), researchers can accurately predict hydrogen storage performance and design optimal porous structures with the ideal balance of micropores and ultra-micropores. Additionally, introducing mesopores and macropores can enhance gas flow and adsorption efficiency. Special attention should be given to the synergistic effects at low temperatures and pressures, which could pave the way for the development of next-generation carbon materials with high hydrogen storage capacity and excellent cyclic stability.

5.3. Gas separation

5.3.1. Carbon dioxide adsorption selectivity

Selectivity is essential in carbon capture, as porous carbon adsorbs not only CO_2 but also N_2. Due to N_2's larger molecular size, it rarely enters micropores, resulting in low adsorption [175]. Optimizing pore structure, especially the micropore-to-mesopore ratio, can enhance CO_2 selectivity while suppressing N_2 adsorption, which is crucial for gas separation and carbon capture applications. A DNN model using textural features and adsorption conditions as inputs found that optimal CO_2 and N_2 adsorption occurs when micropore volume reaches 1.0 cm^3/g and S_{BET} reaches 1250 m^2/g [176]. Increasing mesopores and macropores in porous carbon further suppresses N_2 adsorption due to lower adsorption energy and surface area in these larger pores. To achieve high selectivity, it is recommended to operate at lower pressure, which is also a typical multi-objective optimization problem. Another approach uses N_2 adsorption isotherms at 77 K as inputs in CNN models to predict selectivity. By categorizing adsorption isotherms according to pore size and type, it was predicted that CO_2 adsorption could exceed 5 mmol/g under Type IV isotherms, while N_2 adsorption remains minimal at 298 K and 1 bar, demonstrating high selectivity [177]. Continuous micropore-mesopore structures (2–5 nm) could reduce CO_2 and N_2 adsorption selectivity, so porous carbon materials with well-separated micropores (<2 nm) and mesopores (3–7 nm) in their pore size distribution are ideal for achieving efficient carbon capture.

The optimization of carbon capture selectivity will become a critical research direction in the development of porous carbon materials. First, precise control over pore size distribution remains a key challenge. By combining advanced synthesis techniques with state-of-the-art

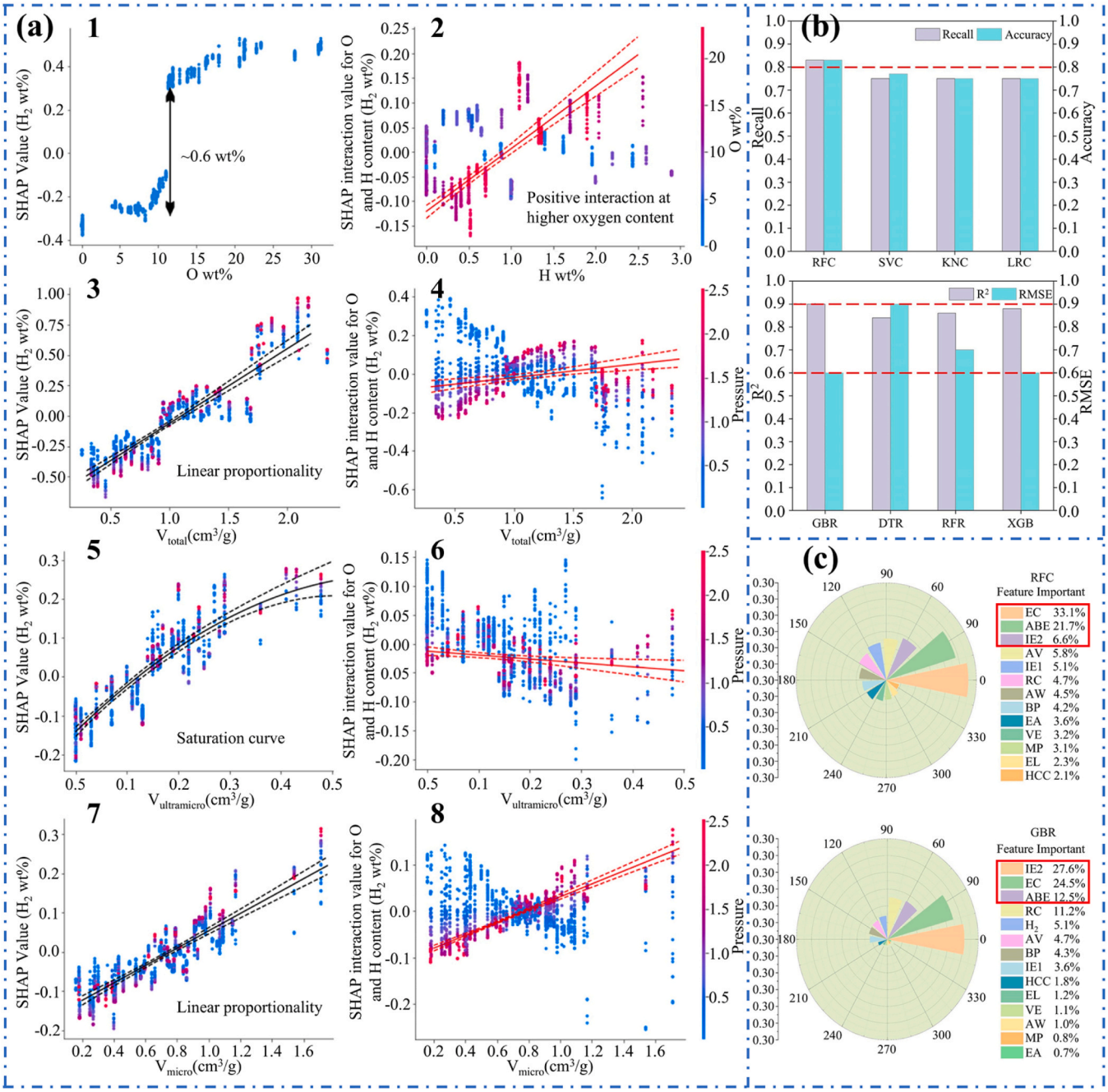


Fig. 13. (a) SHAP values of oxygen content, total pore volume, ultra-micropore volume, and micropore volume on excess hydrogen absorption and the interaction of pressure, oxygen content, etc. Reproduced with permission [167]. copyright 2021, Elsevier. (b) Recall and accuracy of the model for predicting hydrogen storage performance (where C represents the classifier and R represents the regressor). Reproduced with permission [172]. copyright 2025, Elsevier. (c) Feature importance ranking based on simulated data. Reproduced with permission [172]. copyright 2025, Elsevier.

characterization methods, materials with bimodal or multimodal pore size distributions can be designed to ensure high CO_2 adsorption capacity while further improving selectivity. Specifically, under low-pressure conditions, fine-tuning the proportion of micropores (<2 nm) and mesopores (3–7 nm) holds promise for achieving superior CO_2/N_2 separation performance. Additionally, functional modifications will further expand the scope of selectivity control. The introduction of nitrogen-containing groups or other polar functional groups can effectively enhance CO_2 adsorption capacity while suppressing the competitive adsorption of N_2. Future research can focus on the synergistic optimization of pore structures and surface chemistry to achieve efficient separation in multi-component gas systems, paving the way for more advanced and selective carbon capture materials.

5.3.2. *Methane adsorption selectivity*

In applications such as natural gas purification, biogas upgrading, coalbed methane recovery, and industrial hydrogen production, the adsorption selectivity of CH_4/N_2 and CH_4/CO_2 mixtures is a critical factor. Generally, pore size plays a pivotal role in the selective adsorption of gas mixtures under microporous conditions. Confinement within small pores intensifies adsorption competition, enhancing selectivity. For CH_4/N_2 mixtures, the molecular kinetic diameters (3.8 Å for CH_4 and 3.6 Å for N_2) and similar polarities result in typically low selectivity (below 10). However, appropriate design parameters can improve

selectivity, especially in ultra-micropores with high nitrogen content. This is attributed to weak bond interactions between nitrogen dopants and CH_4 molecules, which are more effective in confined spaces for capturing CH_4 [178]. As nitrogen content increases, selectivity for CH_4/N_2 and CH_4/CO_2 improves significantly in small-pore regions [136]. Notably, regions with high selectivity often differ from those with high adsorption capacity. In other words, achieving both high capacity and high selectivity simultaneously is challenging. High adsorption capacity can be enhanced by increasing the contact area between the gas and the adsorbent (i.e., SSA). In contrast, high selectivity requires significant differences in interactions between adsorbate species and the adsorbent. Smaller pore sizes and higher nitrogen content are more favorable for achieving high selectivity.

In gas separation, ML models predict the performance of carbon-based molecular sieve (CMS) membranes. Studies using SVM with an RBF kernel show that gas permeability largely depends on the material's structural characteristics rather than the properties of the gas molecules themselves [179]. ML enables researchers to design optimized pore structures, greatly improving CMS membrane gas separation efficiency. Key factors influencing gas separation performance (such as residual

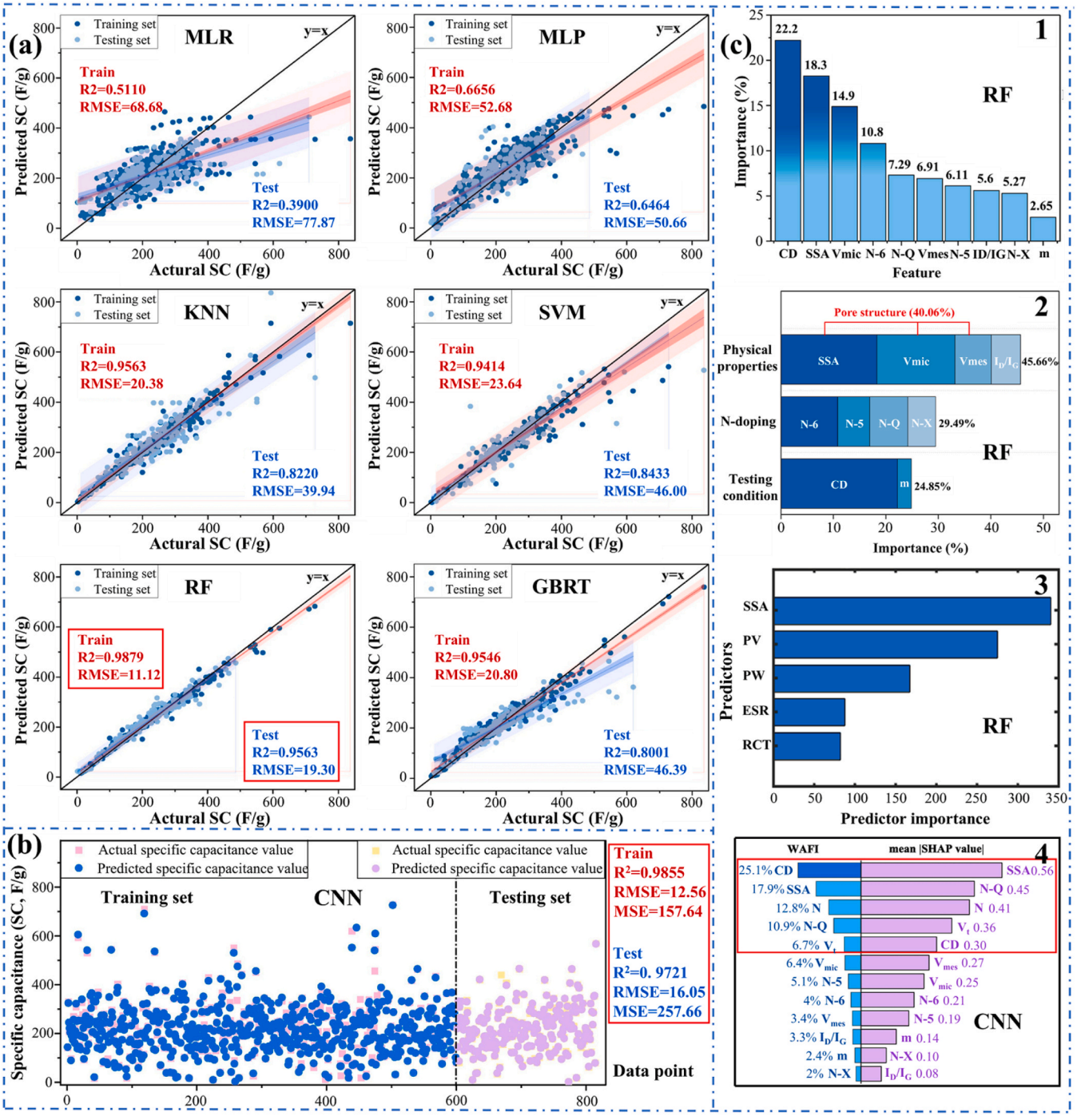


Fig. 14. (a) Accuracy of several models for predicting supercapacitor performance. Reproduced with permission [186]. copyright 2024, Elsevier. (b) Accuracy of CNN for predicting supercapacitor performance. Reproduced with permission [196]. copyright 2024, American Chemical Society. (c) Ranking of feature importance for predicting supercapacitor performance. c_1 and c_2, Reproduced with permission [186]. copyright 2024, Elsevier. c_3, Reproduced with permission [192]. copyright 2024, Institute of Physics. c_4, Reproduced with permission [196]. copyright 2024, Elsevier.

carbon content, free volume fraction of the precursor, and kinetic diameter ratio of gases) can be precisely predicted and optimized using ML models. While these achievements demonstrate ML's potential, its applications in gas separation are still in the early stages, with much work needed to develop more generalized and robust models.

5.4. Energy storage and conversion

The energy storage performance of porous carbon materials is influenced by various factors, and traditional theoretical models often struggle to accurately describe the complex relationships involved. Machine learning models, particularly deep learning approaches, can effectively capture the intricate nonlinear relationships between material structure and performance through training on large sets of experimental and simulation data, enabling high-precision and rapid performance predictions [180]. In the application of porous carbon materials for energy storage, ML significantly enhances the efficiency and accuracy of material design, performance prediction, process optimization, and lifespan analysis [181].

5.4.1. Supercapacitor performance prediction

The performance prediction of carbon-based supercapacitors is influenced by various factors, including material type, SSA, conductivity, pore size distribution, and the conductivity and thermal stability of the electrode system [182,183]. In early studies, methods like LR, SVM, MLP, and DT were commonly used to predict capacitance values, with the overall performance ranking as DT > MLP > SVM > LR, where DT models excelled in capturing complex feature relationships [184,185]. Later, Liu et al. [186,187] tested models including MLR, MLP, KNN, SVM, LGBM [188], RF, and GBDT, using features such as pore size distribution and nitrogen functional group content. They found RF and CNN models to deliver the best performance (Fig. 14a, b). Feature importance rankings for different models in supercapacitor performance prediction yielded similar results [116]. Key factors influencing capacitance values include SSA, micropore volume, total pore volume, and current density (Fig. 14c_1, c_2). For nitrogen functional group distribution, pyridinic-N (N-6) and quaternary-N (N-Q) were the most significant contributors to capacitance, followed by pyrrolic-N (N-5) and N-oxide (N-X) (Fig. 14c_1, c_4). To improve prediction accuracy, ensemble learning methods have been increasingly applied in supercapacitor performance prediction [189]. Ensemble learning integrates the strengths of multiple models to capture the complexities of the input feature space. Recently, ensemble models such as XGB have shown outstanding performance in this field. By optimizing combinations and ratios of various features, studies have identified that S_{mic}/SSA, SSA, and pore size make the largest contributions to capacitance among structural characteristics, providing a promising direction for further capacitance optimization [190].

In recent years, researchers have considered additional factors in supercapacitor performance prediction, such as potential window, electrode configuration, and element doping. In this multi-feature integration approach, machine learning algorithms like ANN and Gaussian Processes have shown strong predictive capabilities, particularly with large datasets, as neural networks can capture complex relationships between conditions and performance [220–222]. For example, Tawfik et al. [191] used an ANN model combined with the SHAP framework to analyze input features and found that SSA, electrode configuration system, and pore volume were most correlated with capacitance, followed by ESR, RCT, and potential window (Fig. 14c_3). To further improve prediction accuracy, they integrated fuzzy logic into the ANN, creating an Adaptive Neuro-Fuzzy Inference System (ANFIS) [192]. However, this model did not include nitrogen/oxygen doping as a feature, so Wang et al. [193] extended the ANN model using data from acidic electrolytes to predict the specific capacitance of N/O co-doped carbon materials. They found that appropriate nitrogen and oxygen doping significantly enhanced the specific capacitance, with the maximum specific surface area of N/O-doped carbon reaching 1502 m^2/g. When the O content was 20 wt% and the N content was 0.5 wt%, the specific capacitance of the N/O co-doped carbon increased from 570 F/g to 611 F/g. Additionally, Golze et al. [194] used the KRR model driven by DFT and GW data to predict XPS spectra and obtained a tool for obtaining XPS of carbon materials in batches. Some scholars later noted that using intermediate variables (e.g., SSA, pore volume) as inputs could introduce noise. Thus, Zhao et al. [192] used initial variables (elemental analysis, industrial analysis, structural composition, activation conditions, and current density) as inputs, applying an LGB model with mutual exclusivity feature binding and gradient unilateral sampling algorithms to prevent overfitting, achieving strong predictive outcomes. Pearson correlation analysis revealed that retaining all initial feature combinations yielded optimal predictive performance. For biomass-based supercapacitors, industrial analysis and structural composition had a greater impact on the model than elemental analysis, with corn stalk-based supercapacitors performing best among agricultural waste biomass. This is likely due to the corn stalk's rich hemicellulose content, whose loose molecular structure and diverse side chains produce numerous oxidation products during carbonization [196]. This has also been demonstrated experimentally. The porous carbon electrode prepared by using $CO(NO_3)_2{\cdot}6\ H_2O$ modified corn straw achieved a maximum specific capacitance of 732.6 F/g and retained 96 % of its capacitance after 5000 cycles, which is at a relatively high level among biomass-based supercapacitors [197].

Traditionally, it is believed that increasing the SSA of electrodes enhances capacitance, with micropores (<2 nm) considered ideal due to their higher SSA compared to mesopores (2–50 nm) or macropores (>50 nm) [223–225]. However, experimental studies have shown that electrodes containing only micropores often exhibit lower power density and suboptimal capacitance performance. This is because micropores increase ionic diffusion resistance, resulting in decreased power density [226,227]. Additionally, characterizing pore size and shape distribution in porous carbon materials experimentally presents significant challenges, especially when micropore sizes are close to ion dimensions, leading to inconsistent findings across studies [228,229]. Thus, further research on pore shape and irregularity in porous carbon materials is a promising direction for future exploration.

As the methods for predicting the performance of carbon-based supercapacitors continue to evolve, future research can delve deeper into several key directions. First, developing more accurate multi-physics coupled prediction models that integrate the electrochemical performance of materials with real-world operating conditions will provide valuable guidance for designing supercapacitors with high stability and long lifespans. Second, exploring the effects of pore shape and distribution on ion transport and capacitance performance can uncover the fundamental mechanisms by which micropores and mesopores contribute to capacitance. Third, expanding performance prediction to multiple dimensions (such as simultaneously predicting capacitance, cycle stability, and power density) will enable the construction of a more comprehensive performance optimization framework. At the same time, addressing the challenges of large-scale production and cost reduction for carbon-based supercapacitors in practical applications remains a critical task. For instance, in the development of biomass-based materials, improving material uniformity and industrial adaptability will significantly advance this field. Moreover, as the significant impact of nitrogen and oxygen doping on performance enhancement becomes increasingly clear, further research into the mechanisms and optimal strategies for element doping could provide new theoretical insights for developing supercapacitors with higher energy and power densities. In summary, by combining experimental and theoretical approaches and integrating algorithmic advancements with material innovations, the energy storage potential of carbon-based supercapacitors will continue to be unlocked, contributing to the realization of a green energy future and a low-carbon society.

5.4.2. Supercapacitor life prediction

The safe operational lifespan of supercapacitors is a critical factor in ensuring system safety. By predicting supercapacitor aging, aging components can be replaced proactively, thus ensuring device reliability and safety [198]. The rapid development of ML offers an effective solution for predicting the remaining useful life (RUL) of supercapacitors. Early studies used life cycle-based neural network models, relying on short charge-discharge curves to predict lifespan. However, due to the limited information in these curves, prediction accuracy was insufficient [199]. To improve RUL prediction, subsequent studies introduced fuzzy neural network models based on time series, incorporating time series, resistance, and capacitance data, which reduced RMSE to 0.01–0.03, achieving more accurate lifespan predictions [200].

With the advancement of CNN and RNN technologies, more studies have applied these models for supercapacitor lifespan prediction. Currently, RUL prediction relies on charge-discharge data collected over continuous cycles under varying voltage, current, and temperature conditions. Given the presence of time-series features in this data, RNNs and their derivatives are well-suited for such applications. For example, Zhou et al. [201] used a Gated Recurrent Unit (GRU) RNN model to analyze the effects of temperature and pressure on supercapacitor aging, providing insights into degradation mechanisms. To further enhance accuracy, Liu et al. [202] employed a stacked bidirectional RNN model, which integrates forward and backward data inputs to account for various factors, offering greater accuracy than a conventional RNN. CNN, known for its excellent feature extraction capabilities in handling images and high-dimensional data, has also been applied to RUL prediction. CNN can autonomously learn critical input features, avoiding human biases. Liu et al. [203] developed a time series-based CNN model, also known as a Temporal Convolutional Network (TCN), which exhibited greater robustness and accuracy than models like back-propagation neural networks, LSTM networks, and GRU. Overall, TCN has shown superior memory functions and feature-capturing abilities in RUL prediction, providing strong support for accurate lifespan prediction of supercapacitors [204]. However, the data-intensive nature of neural networks poses a significant challenge in the context of porous carbon supercapacitors, where data scarcity often limits predictive performance. To address this challenge, some researchers proposed dividing small datasets into subsets of different sizes, thus generating training subsets of various dimensions [205]. The highest-scoring partition is then selected to reduce the accuracy bias caused by data scarcity [206].

Compared to standard RNNs, LSTM, GRU, and TCN each exhibit unique characteristics and advantages. LSTM incorporates a gated mechanism with forget, input, and output gates, enabling it to better capture long-term dependencies in complex time-series data. This makes it particularly well-suited for tasks requiring robust memory retention of long-term information (Fig. 15c). GRU simplifies this structure by using reset and update gates, reducing the number of parameters while maintaining the ability to handle both short- and long-term dependencies. This results in faster training and higher computational efficiency, making GRU ideal for resource-constrained environments and scenarios requiring rapid computation (Fig. 15e) [207]. In contrast, TCN leverages causal and dilated convolutions, offering superior parallelization and effectively capturing long-term contextual information in sequential data while avoiding delays associated with recurrent computations (Fig. 15b). Furthermore, TCN provides greater flexibility and stability by adjusting the dilation factor to expand the receptive field. It

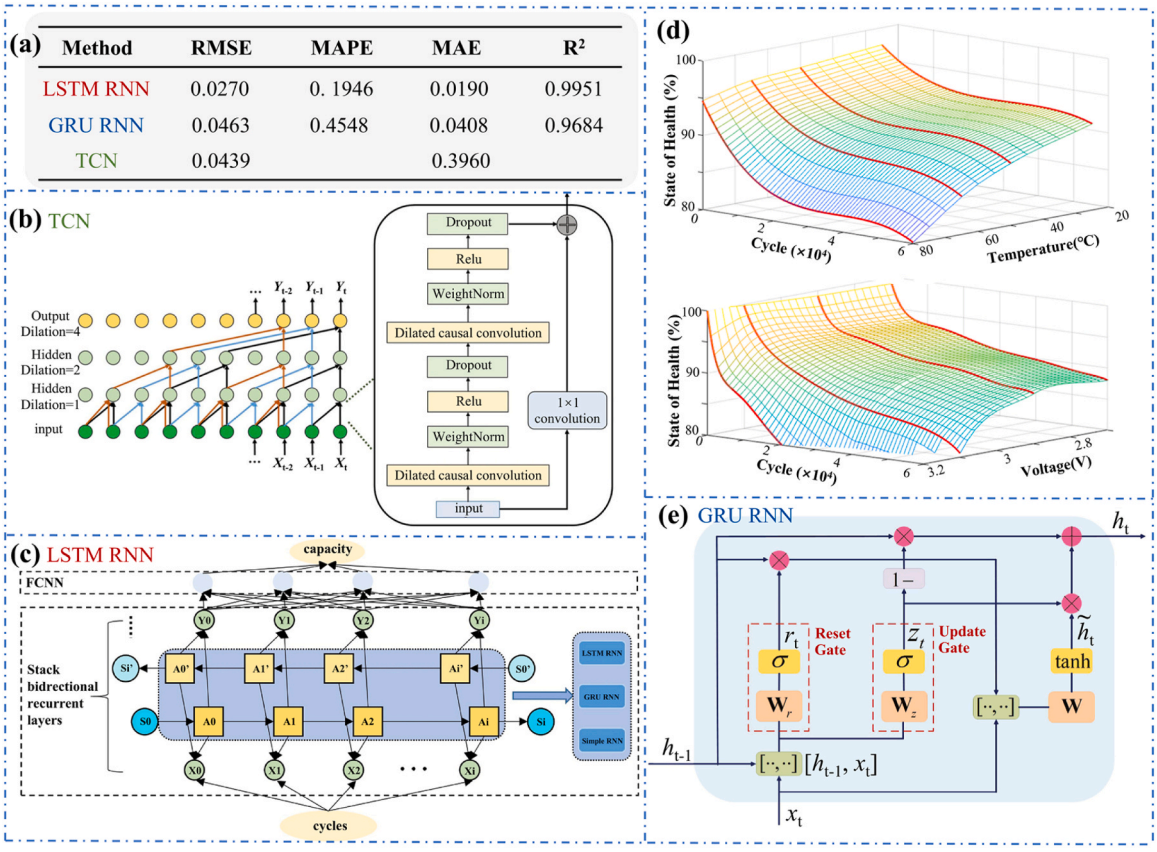


Fig. 15. (a) Accuracy of several supercapacitor life prediction models with time series. Reproduced with permission [202] copyright 2021, John Wiley & Sons. (b) Schematic diagram of the structure of the TCN model. Reproduced with permission [203]. copyright 2022, AIP Publishing. (c) Schematic diagram of the structure of the LSTM RNN model. Reproduced with permission [202] copyright 2021, John Wiley & Sons. (d) The influence of life prediction features. Reproduced with permission [201]. copyright 2019, Elsevier. (e) Schematic diagram of the structure of the GRU RNN model. Reproduced with permission [208]. copyright 2020, Elsevier.

excels at handling large-scale sequence data efficiently and demonstrates better performance in mitigating gradient vanishing or explosion issues. These unique features allow each model to shine in specific task scenarios.

Regarding the factors affecting supercapacitor lifespan, voltage, current, and temperature are significant determinants. Lifespan degradation accelerates notably when temperatures exceed 60℃ or voltages surpass 2.7 V (Fig. 15d). Because there is a clear curve relationship between charge-discharge data over cycles and lifespan, which is relatively simple in complexity, time-series models such as LSTM RNN, GRU RNN, and TCN achieve high prediction accuracy, often exceeding an R^2 of 99.5 % (Fig. 15a). However, this application is considered a straightforward example of lifespan prediction. Future research should incorporate dynamic variables such as ambient temperature, humidity, and charge-discharge rates to develop more precise lifespan prediction frameworks. Real-time lifespan monitoring systems based on edge computing could be developed to provide proactive maintenance and system optimization. Cross-disciplinary collaboration integrating materials science with data science will further enhance our understanding of supercapacitor aging mechanisms, improve feature selection strategies, and elevate the physical interpretability of models. In critical scenarios where supercapacitors operate beyond rated parameters, such as under high temperature, overvoltage, or overcurrent conditions, lifespan degradation often exhibits nonlinear and complex behaviors. This places higher demands on the predictive capabilities of models. Traditional time-series models may struggle to accurately capture sudden failure characteristics in such cases. To address this, advanced methods like reinforcement learning, transfer learning, and physics-informed machine learning models should be explored. These approaches could be combined with finite element simulations (FEM) and multi-scale modeling (MSM) to create more robust predictive frameworks. Additionally, incorporating anomalous parameters representing abrupt changes into the models can help assess their ability to detect abnormal behavior and adapt to extreme conditions. This is crucial for developing reliable lifespan prediction systems capable of handling real-world challenges and ensuring the safe operation of supercapacitors under diverse conditions.

5.4.3. Batteries and catalysis

The application of ML in porous carbon materials is continuously expanding, covering traditional areas like energy storage, adsorption, and supercapacitors, and advancing into cutting-edge fields such as catalysis, gas separation, pollutant treatment, and electrochemical sensors. By incorporating ML techniques, these areas have seen significant improvements in material design efficiency, performance optimization, and application potential [209]. However, the use of ML in these fields is still in its infancy, with relatively few studies exploring its full potential. There remains vast untapped space for future development, particularly in addressing complex challenges and advancing the integration of ML with experimental and theoretical approaches.

In the battery field, ML has been applied to predict sodium storage performance in sodium-ion batteries with porous carbon materials, identifying critical structural parameters in the sodium storage process, such as material crystallinity. This information is particularly valuable for understanding sodium storage mechanisms and predicting performance in non-graphitic carbon materials [210]. Despite these advances, there are still limited studies focusing on broader applications of ML in battery systems, signaling an opportunity for further exploration.

In the field of catalysis, ML is employed to predict the performance of porous carbon materials as catalyst supports and to optimize their design processes, enhancing catalytic activity, selectivity, and stability [211]. For example, ML leverages large-scale data mining and model training to optimize structural characteristics of catalytic active sites, facilitating the design of highly efficient carbon-based catalysts. Wang et al. [212] utilized ML-driven MD simulations at the atomic scale, incorporating van der Waals interactions to retrain the Gaussian approximation potential for carbon materials. Their study on the structure and mechanical properties of nanoporous carbon materials revealed that pore size increases as the density of the carbon material decreases, with the material exhibiting isotropic elasticity that diminishes with increasing density. The intrinsic activity and selectivity of catalysts are primarily determined by the chemical characteristics of functional molecules and the active sites formed after loading, while being synergistically regulated by the textural features of support materials such as SSA, pore size distribution, and surface functional groups [213,214]. The topological structure and defect density of carbon substrates not only influence the loading efficiency and dispersion uniformity of functional molecules but also significantly modulate the catalytic process through electronic effects [215]. Therefore, establishing structure-activity relationship models based on materials genomics, combined with machine learning to predict the evolution of carbon support textural features, and employing density functional theory to elucidate electronic structure changes before and after loading, are of paramount importance for the rational design of high-performance catalysts. Although current ML research has predominantly focused on predicting adsorption energy, catalytic activity, and strength, investigations into catalyst mechanisms, structure-performance relationships, and long-term stability remain relatively limited, presenting promising directions for future exploration.

In summary, although ML has achieved notable progress in the design and optimization of porous carbon materials, its application is still relatively new and underexplored. Currently, research on ML-assisted carbon materials in fields such as batteries, catalysis, and pollutant treatment is still in its early stages. Expanding ML techniques into more complex and interdisciplinary challenges offers tremendous opportunities to revolutionize the field and unlock the full potential of porous carbon materials.

5.5. Pollutant adsorption and treatment

Machine learning has demonstrated remarkable predictive capabilities in the field of pollutant treatment. In a groundbreaking study, Kowalczyk et al. [230] developed an integrated model using Boltzmann-weighted RMSLE scoring combined with multiple regression algorithms, which successfully predicted the paracetamol removal efficiency from aqueous solutions using just three key structural parameters of porous carbon materials: total surface area, ultra-micropore surface area, and mesopore surface area. This breakthrough confirms that ML techniques can effectively predict material performance in pollutant removal based on limited structural features, significantly reducing experimental workload and costs. For emerging antibiotic pollutants, the RF model exhibited outstanding performance in predicting the adsorption capacity of tetracycline (TC) and sulfamethoxazole (SMX) on carbon-based materials, achieving prediction accuracies of R^2= 0.894 and R^2= 0.909, respectively [216]. The RF model not only outperformed GBT and ANN but also demonstrated superior generalization capability compared to traditional isotherm models. Comprehensive feature importance analysis revealed that SSA was the most critical factor, contributing 45 % and 24 % to SMX and TC adsorption, respectively, while solution pH and initial pollutant-to-adsorbent ratio also played significant roles. These three key parameters collectively explained 74 % and 80 % of the variance in TC and SMX adsorption behaviors, with the differences primarily attributed to SMX's smaller molecular size and distinct functional group characteristics compared to TC. Research has shown that porous carbon materials not only exhibit excellent removal efficiency and selective adsorption properties for pharmaceutical pollutants like paracetamol [217,218], but certain modified carbon materials also demonstrate remarkable photocatalytic degradation capabilities for organic dye pollutants [219]. Machine learning techniques can accurately analyze these complex adsorption behaviors and performance differences, providing crucial guidance for the optimized design of carbon-based adsorbents. Notably, ML models show

tremendous potential in predicting carbon materials' performance in adsorbing and photo-catalytically degrading various pollutants, including antibiotics and dyes. By incorporating more intrinsic material properties and environmental parameters, these models could be further extended to practical applications in complex multi-pollutant treatment systems.

6. Future research prospects

6.1. Looking forward to new automated research systems

As early as 2010, researchers had begun exploring the use of robotics to replace traditional experimental processes in materials design [232]. These early systems were predominantly semi-automated material synthesis platforms, where operators set parameters and liquid-handling robots performed the experimental procedures (Fig. 16a). In recent years, with the rapid advancement of ML algorithms, decision-making algorithms for controlling robots in materials design have also seen significant progress. The integration of AI and robotic collaboration has enabled the full automation of experimental processes, wherein AI not only controls the entire experiment but also adjusts experimental conditions based on feedback from material characterization data to optimize material performance (Fig. 16b). For instance, an Autonomous Laboratory combining robotics with ML algorithms has successfully realized fully automated material synthesis, encompassing a closed-loop optimization process from experimental planning to data analysis. This system significantly enhances the efficiency of materials development and applies to various materials, including porous carbon, for optimization and performance prediction [233]. Furthermore, mobile robots have been deployed in automated chemical synthesis laboratories, where they operate experimental equipment to complete complex chemical reactions while leveraging autonomous navigation systems to improve workflow flexibility and efficiency [234]. In addition, AI-driven "robot chemists" not only autonomously perform experimental operations but also optimize reaction conditions based on feedback from high-throughput data, thereby maximizing material performance. These systems have demonstrated tremendous potential in the fields of material synthesis and optimization [235]. In the realm of porous carbon materials, researchers have implemented feedback mechanisms to achieve fully automated processes for the synthesis and optimization of carbon nanotubes. These systems enable the design, execution, and analysis of experiments at speeds several orders of magnitude faster than conventional methods [236,237]. Although the direct application of such technologies to porous carbon materials design remains limited, these advancements hold significant promise for the future optimization and improvement of porous carbon materials. This technology not only accelerates the materials design process but also generates large volumes of experimental data, paving the way for breakthroughs in materials research.

Based on these experimental outcomes, ML models are employed to predict material properties, growth rates, and diameters, etc. Furthermore, using expected improvement-based decision strategies, ML planning models can automatically fine-tune experimental conditions (such as catalyst type, temperature, and reaction time), reducing carbon nanotube diameter control error to about 0.01 nm, surpassing traditional methods in both efficiency and accuracy [238]. By employing trial-and-error in virtual experiments, these models identify optimal parameter combinations, ensuring that actual experiments achieve desired results. The primary challenge in automating ML-assisted porous carbon synthesis lies in the complexity of automating the design steps, particularly with complex precursors. In the future, similar autonomous research systems could be applied across other material synthesis fields, eliminating human influence and significantly enhancing efficiency.

ML can simplify the synthesis process of porous carbon materials by enabling efficient experimental design and accurate prediction models, thus freeing researchers from labor-intensive manual operations and allowing them to focus more on foundational research to explore the microstructure and properties of porous carbon materials. For example, the use of ML techniques can reduce the time required for experimental optimization by up to 50 % and reduce the number of experiments by up to 40 % [239]. Additionally, ML models can rapidly identify materials with potential applications, accelerating their deployment in fields such as energy storage, catalysis, and environmental remediation, further driving technological innovation.

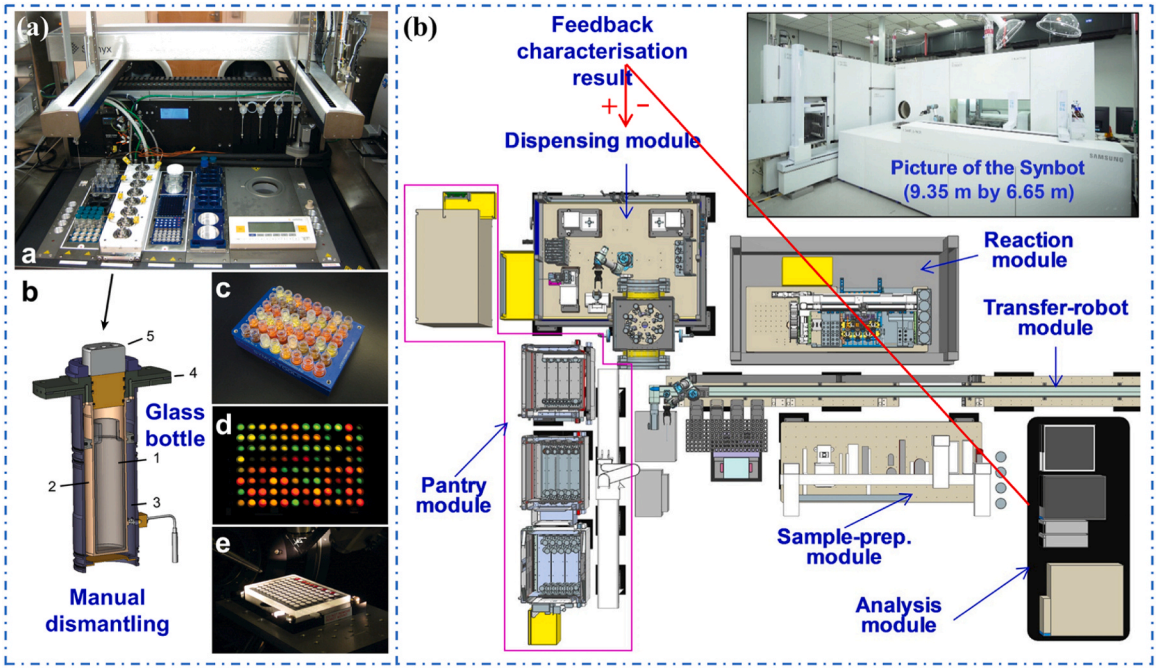


Fig. 16. (a) Semi-automated material synthesis platform. Reproduced with permission [237]. copyright 2016, Springer Nature. (b) Fully automated material synthesis platform. Reproduced with permission [235]. copyright 2023, by T. Ha.

6.2. Looking forward to new research directions

In the design of porous carbon materials, research remains relatively limited, with existing studies primarily emphasizing the surface area and pore volume of biomass-based raw materials. However, abundant carbon-based resources, such as coal, tar pitch, gasification slag, coke, fly ash, and residual carbon, have not been thoroughly investigated. This lack of exploration has resulted in ML studies on carbon material design failing to incorporate critical features related to raw material characteristics, leading to potentially incomplete prediction results.

In terms of the application of porous carbon materials, as summarized in Table 7, current research mainly focuses on carbon capture, energy storage, and supercapacitors. However, many potential ML applications, such as catalyst supports, water treatment, and sensors, remain underexplored, leaving substantial room for future investigation. Even within the existing research directions, there is a notable gap in ML studies addressing the regeneration performance, analytical characteristics, and functional group effects of porous carbon materials used for carbon capture and energy storage. Integrating ML into these areas represents a significant opportunity to advance the field of porous carbon materials. Furthermore, future studies should consider incorporating a broader range of design features, such as pore shapes and a more diverse selection of raw materials. This approach could further drive the development of porous carbon materials, opening new avenues for innovative applications and enhancing the effectiveness of ML-driven material design.

6.3. Looking forward to new MOF carbonization methods

MOFs are a class of rapidly developing functional materials that have gained significant attention over the past decade. Some MOFs exhibit rich pore structures and ultra-high SSA, demonstrating remarkable potential for hydrogen storage. For instance, MOF-210 boasts an SSA exceeding 6000 m^2/g [240]. Additionally, the metal sites on MOF surfaces can further enhance their capabilities in carbon capture and hydrogen storage. However, MOF still faces significant challenges in terms of thermal stability, physicochemical stability, and structural longevity. Studies have shown that the structure, porosity, and hydrogen storage performance of certain MOF cannot be maintained over extended periods [241–243] (under the same storage conditions, carbon materials retain stable hydrogen storage performance and structural integrity after one year, whereas MOF completely lose their hydrogen storage capacity even when carefully preserved for the same duration). This severely limits their industrial applicability. In recent years, researchers have focused on carbonizing high-surface-area MOF into porous carbon materials. However, experiments reveal that the specific surface area of high-surface-area MOF decreases by approximately 40 % during carbonization due to structural collapse and pore blockage [244]. Therefore, a key challenge lies in maximizing the retention of the original carbon framework during pyrolysis while avoiding pore deformation and collapse, which is crucial for producing high-performance porous carbon materials. As shown in Fig. 17, researchers aim to achieve rapid decomposition of tar into small molecules and their swift diffusion during carbonization, while minimizing thermal expansion stress. Nevertheless, the fundamental mechanism behind the rapid decline in specific surface area during MOF carbonization remains poorly understood.

The rapid advancement of ML in recent years has brought new hope to this field. ML has demonstrated significant potential in analyzing MOF carbonization processes. Its image recognition technology can accurately capture the dynamic evolution of MOF molecular frameworks during carbonization through high-dimensional data analysis, revealing key nodes where the detachment of volatile groups such as H, O, and N leads to structural collapse. This provides a theoretical foundation for identifying the most stable framework configurations during carbonization. Furthermore, ML employs data-driven approaches to deeply investigate the multi-scale mechanisms of carbonization, integrating high-throughput computational and experimental data to establish complex mappings between external conditions (e.g., temperature, pressure, and atmosphere) and MOF structural evolution. This

Table 7
Summary of ML predictions for porous carbon performance in different application scenarios.

CO_2 Capture		CH_4 Storage		H_2 Storage		Supercapacitors		Catalyst carrier		Water treatment		Sensor	
Specific surface area	✓	Specific surface area	✓	Specific surface area	✓	Specific surface area	✓	Specific surface area	✗	Specific surface area	✗	Specific surface area	✗
Pore volume	✓	Pore volume	✓	Pore volume	✓	Pore volume	✗	Pore volume	✗	Pore volume	✗	Pore volume	✗
Pore diameter	✓	Pore diameter	✗	Pore diameter	✗	Pore diameter	✗	Pore diameter	✗	Pore diameter	✗	Pore diameter	✗
Pore irregularity	✗	Pore irregularity	✗	Pore irregularity	✗	Pore irregularity	✗	Pore irregularity	✗	Pore irregularity	✗	Pore irregularity	✗
Adsorption capacity	✓	Adsorption capacity	✓	Adsorption capacity	✓	Conductivity	✗	Adsorption capacity	✗	Adsorption capacity	✓	Conductivity	✗
Adsorption energy	✓	Adsorption energy	✗	Adsorption energy	✗	Specific Capacitance	✓	Adsorption energy	✓	Adsorption energy	✗	Response time	✗
Adsorption rate	✗	Adsorption rate	✗	Adsorption rate	✗	Energy density	✓	Thermal stability	✗	pH stability	✗	Recovery time	✗
Selectivity	✓	Selectivity	✓	Selectivity	✗	Power density	✓	Active	✓	Selectivity	✓	Selectivity	✗
Desorption performance	✗	Desorption performance	✗	Desorption performance	✗	Bulk density	✗	Regeneration performance	✓	Regeneration performance	✗	Sensitivity	✗
Regeneration performance	✗	Regeneration performance	✗	Regeneration performance	✗	Cycle life	✓	Antitoxicity	✗	Anti-interference	✗	Environmental stability	✗
Functional groups	✗	Chemical modification	✗	Chemical modification	✗	Internal resistance	✗	Active site	✓	Functional groups	✗	Functional groups	✗

Note: A check mark indicates that ML has already been used for predictions, while a cross mark indicates that ML has not yet been used for predictions.

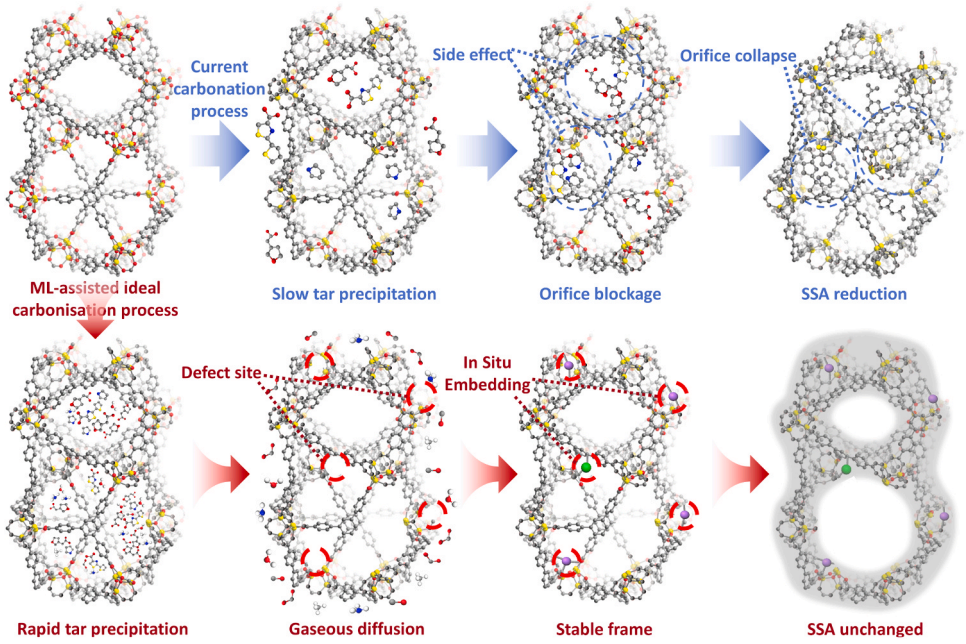


Fig. 17. Bottlenecks in the MOF carbonization process and the ideal carbonization process assisted by ML.

enables the prediction of pore structure, SSA, and stability of carbonized MOF under various conditions. RL-based optimization algorithms can automatically explore optimal carbonization parameters, such as heating rate, pyrolysis temperature, and duration, to minimize structural collapse and pore blockage. Additionally, ML can screen candidate materials with high stability during carbonization by analyzing the chemical composition and topological structure of MOF precursors, providing theoretical guidance for experimental design. Given the vast diversity of MOF and the large number of atoms within their frameworks, traditional molecular dynamics simulations face challenges of high computational complexity and time consumption. In contrast, ML significantly reduces computational costs by constructing efficient surrogate models while maintaining high-precision predictions of complex physicochemical processes. This offers a new research paradigm for optimizing MOF carbonization pathways and designing high-performance porous carbon materials, driving the transition from experience-driven to data-driven approaches in this field. These research directions await further exploration by scholars.

6.4. *Looking forward to new ML models*

In the performance prediction of porous carbon materials, RF and deep learning models exhibit high accuracy, which can be attributed to several factors. RF models effectively handle complex nonlinear relationships and reduce overfitting risks through the integration of multiple decision trees, making them particularly well-suited for analyzing multidimensional and multi-feature experimental data [245–249]. Meanwhile, deep learning models, such as DNN and CNN, excel at capturing deep features and patterns through multi-layer nonlinear mapping, demonstrating significant advantages in processing complex data, such as pore structure characteristics and adsorption performance. Additionally, the robust self-learning capability of deep learning models allows them to optimize feature weights during large-scale data training, thereby enhancing prediction performance [250–254].

However, the effectiveness of these models is influenced by data quality and feature selection. For instance, RF models are relatively sensitive to the independence of input features, while deep learning models rely heavily on the scale and diversity of the data [255–258]. Future improvements to these models can focus on several directions: Integrating multimodal learning to combine experimental data with simulation data and characterization images to enhance model comprehensiveness; Optimizing feature engineering by leveraging explainable algorithms like SHAP to identify key features and using automated feature generation methods to improve input quality; developing more efficient hybrid models that combine RF with deep learning to leverage their respective strengths and improve prediction accuracy and robustness; and embedding physical constraints, such as employing reinforcement learning to incorporate physical laws into the learning and prediction processes, thereby improving model performance and reliability. These advancements will enable more accurate predictions of the performance of porous carbon materials across various applications, providing stronger theoretical support and practical guidance for material design.

6.5. *Looking forward to new ML features*

Due to the presence of unsaturated chemical bonds, the surface of porous carbon materials forms adsorption fields, where gas adsorption is influenced primarily by van der Waals forces or weak chemical bonds, resulting in positional or non-positional adsorption. The CO_2 adsorption density decreases as pore size increases, mainly because smaller pores provide stronger adsorption potential between the pore walls and CO_2 molecules [259,260]. The more confined gas molecules are within the pore channels, the stronger the interactions with the pore walls, leading to higher adsorption potential [261]. Once gas molecules are captured by the pore walls, the enhanced adsorption potential makes it easier for the molecules to be trapped, thereby increasing adsorption density [262]. In applications such as carbon capture, energy storage, and water treatment, adsorption can be divided into two main types: adsorption on solid surfaces and multilayer compressed aggregates within slit pores [263,264]. Especially in slit pores, gases exhibit a multilayer adsorption phenomenon. During physical or chemical adsorption, gas molecules attach to porous carbon walls through intermolecular forces or weak chemical bonds, gradually forming multilayers as pressure and spatial restrictions increase. As shown in Fig. 18, adsorption processes differ under low-pressure and high-pressure conditions. At low pressure, the mechanism corresponds to volume filling, while at high pressure, it corresponds to layer-by-layer adsorption. Under low pressure, the kinetic energy of gas molecules is relatively low, making them more likely to be confined within small micropores or trapped at surface

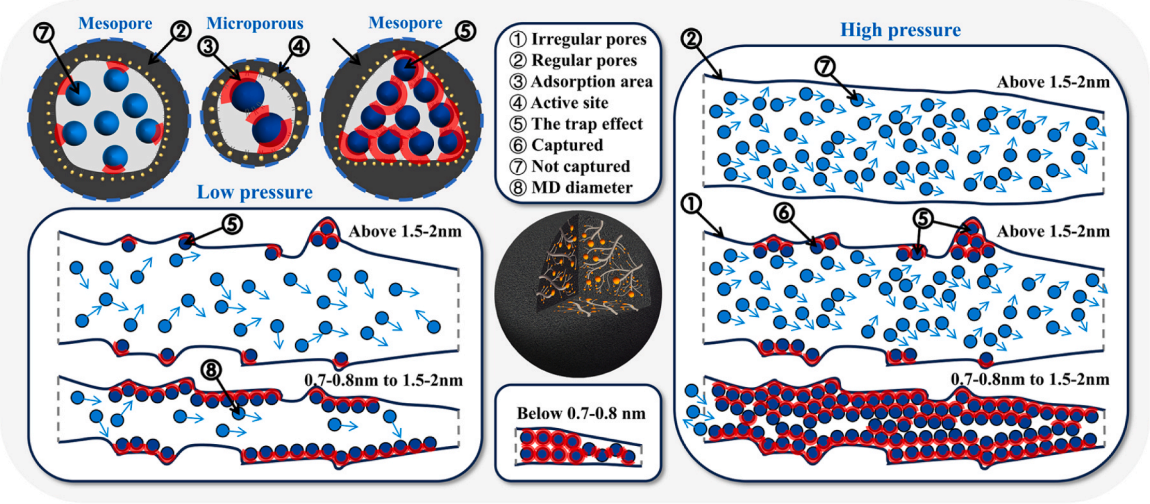


Fig. 18. Schematic illustration of gas molecule adsorption mechanisms in irregular pores under high and low pressure, showing the volumetric filling mechanism at low pressure, the layer adsorption mechanism at high pressure, and the "trap effect" in recessed regions.

depressions. In mesopores, gas molecules first fill the depressions, then rapidly fill the entire pore volume rather than adhering layer by layer [265,266]. At high pressure, as the concentration of gas molecules increases, more molecules overcome the barriers of thermal motion and are captured by the pore walls. After the small micropores and surface depressions are filled, gas molecules begin to stack layer by layer within larger pores, forming multilayer adsorption. At this stage, gas molecules are adsorbed onto the pore walls rather than filling the pore volume [267–269].

Two main approaches enhance the adsorption capacity of porous carbon materials for gases like hydrogen, methane, and carbon dioxide [270]. The first approach involves doping elements or oxidizing porous carbon materials to increase surface adsorption energy, a method well-supported by existing studies approach is to increase the adsorption area, an area with limited research [271]. In Fig. 18, with the coverage of adsorption layers, the pore size progressively decreases, intensifying confinement effects on gas molecules. Under identical pore sizes, pores with higher irregularity can offer a larger adsorption surface area, facilitating the quicker formation of the first adsorption layer and thereby promoting multilayer adsorption. Conversely, regular pores have a weaker adsorption effect and may not form a complete adsorption layer. This highlights the potential of irregular porous structures to enhance adsorption capacity, providing a new research direction for future studies.

In the porous channels of materials, gas molecules are also influenced by van der Waals forces or dipole interactions from surrounding gas molecules. These interactions can include either attractive or repulsive forces, depending on the relative positions and charge distributions of the molecules [272,273]. As gas molecules approach the channels or move closer to the pore walls, they must overcome these opposing forces, including certain repulsive forces, before they can adhere to the walls and be captured [274]. When one or more layers of gas molecules have already been adsorbed within the pores, van der Waals attraction occurs between the already adsorbed molecules and newly incoming gas molecules. At the same time irregular pore shapes can restrict the movement of gas molecules, thus increasing the probability of collision of gas molecules with the wall, leading to easier attractive trapping. During multilayer adsorption, the adsorption surface (comprising both the pore walls and the previously adsorbed molecules) forms an extended adsorption potential well. As new molecules approach, they can "step into" this potential well gradually, without having to overcome a single high-energy barrier as in monolayer adsorption. In other words, for the first adsorption layer, gas molecules must overcome strong repulsive forces between themselves and the pore walls. However, for the second layer, gas molecules only need to overcome weaker repulsive forces between themselves and the first layer of adsorbed molecules. As the number of adsorbed layers increases, the repulsive and attractive forces on the new molecules decrease until an equilibrium state of adsorption and desorption is reached.

As shown in Fig. 18, in recessed areas, gas molecules interact with multiple wall regions around the depression through van der Waals forces or electrostatic interactions. These interactions are additive, resulting in enhanced adsorption capacity. This phenomenon is analogous to a "trap" effect, where molecules entering the recessed areas experience attraction from multiple directions, making them less likely to escape. The "encapsulation" of recessed areas increases the contact area between gas molecules and the surface, thereby strengthening the overall adsorption forces. In contrast, on protruding regions, gas molecules primarily interact with single points on the surface, resulting in limited contact areas and weaker interaction forces. Furthermore, because protruding regions are often exposed to the interior of the pore or the pore entrance, gas molecules are more likely to be carried away by fluid motion or diffusion, making stable adsorption difficult. In practical research, optimizing adsorption performance can be achieved by designing a large number of micropores or recessed structures to enhance adsorption capacity while reducing the proportion of smooth surfaces or protruding regions.

Overall, pore irregularity significantly enhances the adsorption performance of porous carbon materials. In Fig. 19, the formation process of porous carbon materials is highly complex, involving gas penetration, deposition, release, as well as the generation and expansion of pores, all of which are influenced by the composition of raw materials and pyrolysis conditions. However, pore irregularity arises as a result of the combined effects of these processes, and it is further shaped by the properties of the raw materials, the pyrolysis techniques, and the specific pyrolysis methods used. If experimental conditions can be effectively linked to pore irregularity characteristics, this feature could serve as a key indicator of how different pore shapes influence the adsorption capacity of gas molecules. This would provide valuable insights for the integration of ML in the design and application of porous carbon materials.

The raw material composition of porous carbon materials is highly varied, with carbon present in multiple forms. The content and distribution of active and inert components, along with differences in decomposition rates and timing among components, contribute to the diversity of pore shapes. Additionally, different constituents or carbon

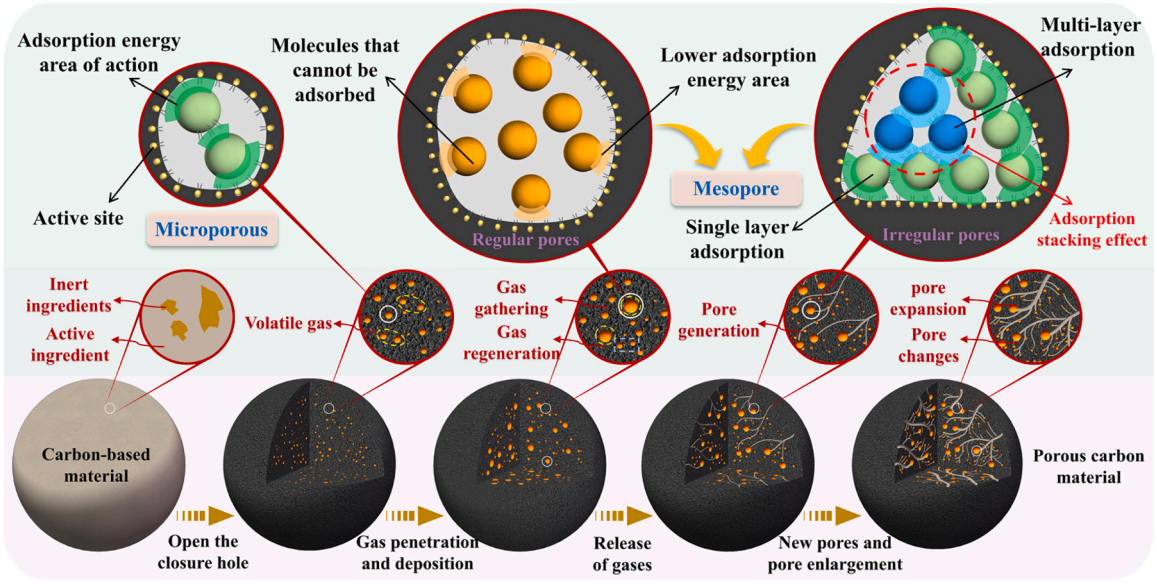


Fig. 19. The process of pore formation and the influence of pore shape on adsorption.

structures exhibit varying hardness levels, resulting in uneven distribution of internal stress. The type and amount of activating agents significantly affect the pore shapes in carbon materials. In Fig. 20, the processes of softening, pore collapse, stress release, and re-carbonization of carbon materials vary at different temperatures, further impacting pore irregularity. Thus, key characteristics of raw material composition and pyrolysis conditions play a critical role in determining the complexity of pore shapes. A thorough understanding of these influencing factors can enable more precise control over pore structure, which is essential for optimizing the performance of porous carbon materials in applications like adsorption, energy storage, and catalysis.

Previous studies indicate that pore shapes significantly impact the effectiveness of porous carbon materials in adsorbing gas molecules, solution ions, and macromolecular pollutants. However, due to the physical and chemical randomness of pore irregularity, this factor is often overlooked. ML offers unique advantages in uncovering the patterns in these seemingly random features, making pore shape a key factor for future research. SEM images contain valuable information on the shape characteristics of porous carbon material pores. Scholars like Kurbakov et al. [275] have conducted related research, utilizing SEM images of carbon-supported nanoparticles. They defined custom nanoparticle orderly characteristics and used a CNN model to identify ordered and disordered arrangements of palladium nanoparticles, analyzing defect distribution on carbon material surfaces. This exploration marks the beginning of automated SEM image analysis in carbon materials and materials science. Similarly, pore shape information can also be extracted from SEM images, providing additional data on pore shape features. However, due to considerable noise in SEM images, appropriate denoising is essential before feature extraction. For some carbon materials that are challenging to synthesize, such as vertically aligned carbon nanotubes, obtaining numerous authentic SEM images is difficult. Hajilounezhad et al. [276] proposed generating SEM images through virtual labs based on physical simulations, which, combined with deep learning models, can predict the structure and mechanical properties of carbon nanotubes. This method offers an effective alternative for data-scarce scenarios, broadening the scope of SEM-based pore structure analysis in ML applications for carbon materials.

As a tool, ML has certain limitations. To fully harness its potential, it is recommended to use it in conjunction with other methods, such as experimental techniques or DFT. This approach can help clarify the mechanisms behind material performance improvements and create a synergistic effect by leveraging the strengths of different methods. However, the true value of ML in the synthesis and performance prediction of porous carbon materials lies in uncovering unexpected structural characteristics and effects. It cannot replace researchers' enthusiasm and creativity in exploring the structure-performance relationships of carbon materials. ML serves as a powerful tool to reveal hidden patterns and correlations, but the insights and innovative approaches in material design still heavily rely on human intuition and experimental curiosity. Ultimately, ML complements human ingenuity, enabling more efficient discovery while inspiring deeper exploration into the fundamental mechanisms of porous carbon materials.

6.6. *Perspective*

In this field, the development of high-quality datasets represents

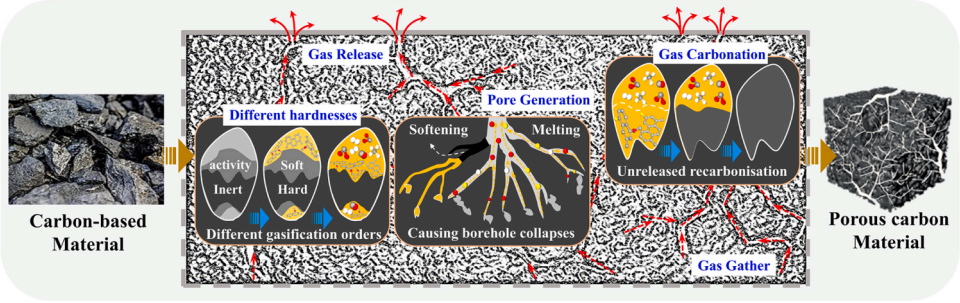


Fig. 20. The complex carbonization and activation process of carbon materials.

both a key research focus and a major bottleneck, as currently available public databases remain extremely limited. This research direction is still in its nascent stage, with numerous subfields awaiting machine learning-assisted advancement. All of which fundamentally depend on specialized dataset support. Therefore, we strongly advocate for academic collaboration to establish standardized, open-access domain-specific databases. Moreover, the machine learning methods and approaches for carbon materials presented in this study are equally applicable to other materials science research domains.

In the design and application of porous carbon materials, the characterization of pore structures and the quantification of irregularity represent two critical challenges. Irregular pores can significantly enhance adsorption performance through additional active sites created by surface concavities. Two primary approaches are available for evaluating pore irregularity: (1) morphological analysis by quantifying the number and depth of concavities in 2D/3D pore structures to establish quantitative metrics, and (2) comprehensive analysis combining SSA and PA. Concurrently, key parameters influencing material performance have been progressively identified in carbon material applications. While existing predictive models still have room for accuracy improvement, the potential gains may be limited. Consequently, research focus is shifting toward automated material optimization design, whose core lies in establishing a closed-loop workflow of "automated characterization → model optimization → feedback adjustment → experimental verification". The efficient coordination of these interconnected processes has emerged as a major technical challenge.

7. Conclusion

In summary, this review provides a comprehensive analysis of ML methods in the synthesis and application of porous carbon materials, with a particular focus on areas like carbon capture, energy storage, and supercapacitors. With efficient predictive models, ML has the potential to streamline the synthesis process of porous carbon materials, freeing scientists from labor-intensive manual experiments. The integration of robotics with ML holds promise for developing autonomous research systems that could significantly accelerate the discovery of new materials. In terms of applications, ML not only enhances the performance of porous carbon materials but also offers synthesis guidance. The continued development of novel porous carbon materials will drive the advancement and application of various emerging technologies. Based on the advances in various algorithms for different applications, integrated models are likely to dominate ML research in porous carbon materials moving forward. This paper also provides an in-depth analysis of feature importance across applications, emphasizing the critical roles of ultra-micropores and some mesopores.

Additionally, while the literature frequently highlights the significant impact of pore shape on adsorption performance, the underlying mechanisms remain underexplored. Currently, ML applications remain concentrated in fields like carbon capture, energy storage, and supercapacitors, with many other applications awaiting to be explored. Although some studies have focused on surface area and pore volume characteristics, they are mostly limited to biomass materials; raw materials such as coal, tar pitch, gasification slag, coke, fly ash, and residual carbon have yet to be deeply investigated. This reveals the vast potential for further exploration of ML in the field of porous carbon materials.

CRediT authorship contribution statement

Luo Kai Hong: Writing – review & editing, Formal analysis. **Cheng Xingxing:** Writing – review & editing, Funding acquisition. **Wang Chuang:** Writing – original draft, Methodology. **Bi Xiaotao:** Writing – review & editing, Data curation. **Ni Meng:** Investigation, Conceptualization. **Wang Zhiqiang:** Supervision, Project administration. **Nandakumar Krishnaswamy:** Investigation, Conceptualization. **Wang Chunbo:** Supervision, Data curation. **Zhang Jiansheng:** Resources, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the Shandong Province Natural Science Foundation (NO. ZR2024ME243 and NO. ZR2024ME120), the Shanxi Key Research and Development Plan (NO. 202201090301002), the Xinjiang Key Research and Development Plan (NO. 2024B04001) and the National Natural Science Foundation of China (Grant No. 52250710681) for the financial support.

Data availability

Data will be made available on request.

References

[1] A. Jaisan, A. Mohan, Y. Lee. Machine learning-enhanced photocatalysis for environmental sustainability: Integration and applications, 2024, 161: 100880. ⟨https://doi.org/10.1016/j.mser.2024.100880⟩.
[2] Z.Q. Yang, B.S. Chen, H.M. Chen, et al., A critical review on machine-learning-assisted screening and design of effective sorbents for carbon dioxide (CO_2) capture, Front. Energy Res. 10 (2023) 2022, https://doi.org/10.3389/fenrg.2022.1043064.
[3] Z.L. Li, L.B. Deng, L.A. Kinloch, et al., Raman spectroscopy of carbon materials and their composites: graphene, nanotubes and fibres, Prog. Mater. Sci. 135 (2023) 101089, https://doi.org/10.1016/j.pmatsci.2023.101089.
[4] S. Jo, K. Lee, Y. Jung, et al., Direct-spun carbon nanotube sheet: a flexible, ultralight, stackable three-dimensional current collector for high-performance lithium-ion batteries, Carbon 219 (2024) 118786, https://doi.org/10.1016/j.carbon.2024.118786.
[5] J.M. Wei, J.L. Lou, W.B. Hu, et al., Superstructured carbon with enhanced kinetics for zinc-air battery and self-powered overall water splitting, Small 20 (2024) 2308956, https://doi.org/10.1002/smll.202308956.
[6] G. Conte, A. Policicchio, M. Idrees, et al., Tuning the ultra-microporosity in hierarchical porous carbons derived from amorphous cellulose towards a sustainable solution for hydrogen storage, Int. J. Hydrog. Energy 50 (2024) 763–773, https://doi.org/10.1016/j.ijhydene.2023.08.128.
[7] R. Muhammad, J. Park, H. Kim, et al., Facile synthesis of ultrahigh-surface-area and hierarchically porous carbon for efficient capture and separation of CO_2 and enhanced CH_4 and H_2 storage applications, Chem. Eng. J. 473 (2023) 145344, https://doi.org/10.1016/j.cej.2023.145344.
[8] T. Lemaoui, T. Eid, A.S. Darwish, et al., Revolutionizing inverse design of ionic liquids through the multi-property prediction of over 300,000 novel variants using ensemble deep learning, Mater. Sci. Eng.: R: Rep. 159 (2024) 100798, https://doi.org/10.1016/j.mser.2024.100798.
[9] W.F. Cai, X.X. Zhu, R. Kumar, et al., Catalytic pyrolysis of biomass waste using montmorillonite-supported ultrafine iron nanoparticles for enhanced bio-oil yield and quality, Green. Energy Resour. 2 (2024) 100085, https://doi.org/10.1016/j.gerr.2024.100085.
[10] X. Bai, X. Zhang, Artificial intelligence-powered materials science, Nano-Micro Lett. 17 (2025) 135, https://doi.org/10.1007/s40820-024-01634-8.
[11] Y.T. Mo, X. Deng, P.X. Liu, et al., Insights into the application of carbon materials in heterojunction solar cells, Mater. Sci. Eng.: R: Rep. 152 (2023) 100711, https://doi.org/10.1016/j.mser.2022.100711.
[12] Y. Wang, Application-oriented design of machine learning paradigms for battery science, npj Comput. Mater. 11 (2025) 89, https://doi.org/10.1038/s41524-025-01575-9.
[13] E.V. Luis, L.M. Claudia, M. Cecilia, et al., Machine learning and materials informatics approaches in the analysis of physical properties of carbon nanotubes: a review, Comput. Mater. Sci. 201 (2022) 110939, https://doi.org/10.1016/j.commatsci.2021.110939.
[14] X.M. Wang, Y.J. Xiao, T. Yang, et al., Quantitative assessment of cement bridges and voids in cement-stabilized permeable base materials using a mask R-CNN-based CT image segmentation strategy, Mater. Des. 241 (2024) 112907, https://doi.org/10.1016/j.matdes.2024.112907.
[15] Z. Yang, Q. Yu, Y. Zhan, et al., Incorporating edge convolution and correlative self-attention into graph neural network for material properties prediction, Mach. Learn.: Sci. Technol. 6 (2025) 015020, https://doi.org/10.1088/2632-2153/ad9fcf.
[16] S.C. Joshi, Knowledge based data boosting exposition on CNT-engineered carbon composites for machine learning, Adv. Compos. Hybrid. Mater. 3 (2020) 354–364, https://doi.org/10.1007/s42114-020-00171-3.

[17] K.J.G. Tun, L. Elizabeth, G. Octavio, Optimal control strategy based on artificial intelligence applied to a continuous dark fermentation reactor for energy recovery from organic wastes, Green. Energy Resour. 3 (2025) 100112, https://doi.org/10.1016/j.gerr.2024.100112.

[18] J. Sang, B. Tang, K.C. Pan, et al., Current status and enhancement strategies for All-Solid-State lithium batteries, Acc. Mater. Res. 4 (6) (2023) 472–483, https://doi.org/10.1021/accountsmr.2c00229.

[19] Y.M. Liu, X.Y. Tan, J. Liang, et al. Machine learning for perovskite solar cells and component materials: Key technologies and prospects. 2023, 33: 2214271. ⟨https://doi.org/10.1002/adfm.202214271⟩.

[20] Y. Ihara, H. Yamagishi, M. Naito, et al., Machine learning of organic solvents reveals an extraordinary axis in Hansen space as indicator of spherical precipitation of polymers, Aggregate 4 (2023) e365, https://doi.org/10.1002/agt2.365.

[21] G. Huang, Y. Guo, Y. Chen, et al., Application of machine learning in material synthesis and property prediction, Materials 16 (2023) 5977, https://doi.org/10.3390/ma16175977.

[22] A. Akinpelu, M. Bhullar, Y. Yao, Discovery of novel materials through machine learning, J. Phys.: Condens. Matter 36 (2024) 453001, https://doi.org/10.1088/1361-648X/ad6bdb.

[23] Z.H. Shen, H.X. Liu, Y. Shen, et al., Machine learning in energy storage materials, Interdiscip. Mater. 1 (2022) 175–195, https://doi.org/10.1002/idm2.12020.

[24] V. Pavlenko, S. Khosravi, S. Żółtowska, et al., A comprehensive review of template-assisted porous carbons: Modern preparation methods and advanced applications, Mater. Sci. Eng.: R: Rep. 149 (2022) 100682, https://doi.org/10.1016/j.mser.2022.100682.

[25] I. Pet, M.N. Sanad, M. Farouz, et al., Review: recent developments in the implementation of activated carbon as heavy metal removal management, Water Conserv. Sci. Eng. 9 (2024) 62, https://doi.org/10.1007/s41101-024-00287-3.

[26] Z.Y. Gan, S. Callegari, S.A. Cobbold, et al., Activation mechanism of PINK1, Nature 602 (2022) 328–335, https://doi.org/10.1038/s41586-021-04340-2.

[27] J.X. Cheng, Z.J. Lu, X.F. Zhao, et al., Green needle coke-derived porous carbon for high-performance symmetric supercapacitor, J. Power Sources 494 (2021) 229770, https://doi.org/10.1016/j.jpowsour.2021.229770.

[28] T. Pei, F. Sun, J.H. Gao, et al., Introducing catalytic gasification into chemical activation for the conversion of natural coal into hierarchically porous carbons with broadened pore size for enhanced supercapacitive utilization, RSC Adv. 8 (2018) 37880–37889, https://doi.org/10.1039/c8ra07308b.

[29] Y. Liu, G.C. Yan, J. Wang, et al., Liquefaction pitch-based porous carbon: preparation and relationship between pore structure and electrochemical properties, Diam. Relat. Mater. 122 (2022) 108824, https://doi.org/10.1016/j.diamond.2022.108824.

[30] G. Li, Z.G. Sun, Y.Y. Zhang, et al., One-step green synthesis of nitrogen and phosphorus co-doped pitch-based porous graphene-like carbon for supercapacitors, J. Porous Mater. 24 (2017) 1689–1696, https://doi.org/10.1007/s10934-017-0411-4.

[31] Y.W. Zhu, Y.J. Wang, T.Y. Wang, et al., One-step preparation of coal-based magnetic activated carbon with hierarchically porous structure and easy magnetic separation capability for adsorption applications, J. Magn. Magn. Mater. 569 (2023) 170480, https://doi.org/10.1016/j.jmmm.2023.170480.

[32] M. Bora, J. Tamuly, S.M. Benoy, et al., Highly scalable and environment-friendly conversion of low-grade coal to activated carbon for use as electrode material in symmetric supercapacitor, Fuel 329 (2022) 125385, https://doi.org/10.1016/j.fuel.2022.125385.

[33] R. Wang, T. Lee, S.K. Rish, et al., One-pot synthesis of N-doped carbon nanosheets from Victorian brown coal with enhanced lithium storage, Fuel Process. Technol. 238 (2022) 107498, https://doi.org/10.1016/j.fuproc.2022.107498.

[34] Z.M. Du, Z.P. Lei, H.L. Yan, et al., HNO_3 pre-oxidation-tuned microstructures of porous carbon derived from high-sulfur coal for enhancing capture and catalytic conversion of polysulfides, Fuel 326 (2022) 125066, https://doi.org/10.1016/j.fuel.2022.125066.

[35] M. Shi, Y.F. Xin, X.X. Chen, et al., Coal-derived porous activated carbon with ultrahigh specific surface area and excellent electrochemical performance for supercapacitors, J. Alloy. Compd. 859 (2021) 157856, https://doi.org/10.1016/j.jallcom.2020.157856.

[36] X. Wang, Z.H. Qin, X.Q. Yang, et al., Synthesis of tremella-like porous carbon with oxygen-containing functional groups for high-performance supercapacitor, Diam. Relat. Mater. 108 (2020) 107995, https://doi.org/10.1016/j.diamond.2020.107995.

[37] Y.T. Xu, X.L. Chai, Characterization of coal gasification slag-based activated carbon and its potential application in lead removal, Environ. Technol. 39 (3) (2018) 382–391, https://doi.org/10.1080/09593330.2017.1301569.

[38] J. Abdulsalam, J. Mulopo, B. Oboirien, et al., Experimental evaluation of activated carbon derived from South Africa discard coal for natural gas storage, Int. J. Coal Sci. Technol. 6 (3) (2019) 459–477, https://doi.org/10.1007/s40789-019-0262-5.

[39] L.L. Sun, Z.B. Zhao, Y. Sun, et al., Activated coal-based graphene with hierarchical porous structures for ultra- high energy density supercapacitors, Diam. Relat. Mater. 106 (2020) 107827, https://doi.org/10.1016/j.diamond.2020.107827.

[40] M. Zhao, X. Bai, X. Fan, et al., Removal behaviors of phenol from aqueous solution using industrial coal sludge-derived porous carbon sorbent, J. Mol. Liq. 385 (2023) 122427, https://doi.org/10.1016/j.molliq.2023.122427.

[41] J.B. Zhang, G.R. Zhang, M. Qi, et al., Co-production of hydrogen-rich gas and porous carbon by partial gasification of coal char, Chem. Pap. 72 (2018) 273–287, https://doi.org/10.1007/s11696-017-0278-5.

[42] Z. Miao, Z.K. Guo, G.F. Qiu, et al., Synthesis of activated carbon from high-ash coal gasification fine slag and their application to CO_2 capture, J. CO_2 Util. 50 (2021) 101585, https://doi.org/10.1016/j.jcou.2021.101585.

[43] J.L. Gonzalez, Multifractal statistic of porous structures simulated through fully penetrable sphere models, Phys. A: Stat. Mech. Appl. 567 (2021) 125695, https://doi.org/10.1016/j.physa.2020.125695.

[44] R. Ding, X.B. Wang, A. Tian, et al., Unlocking new insights for electrocatalyst design: A unique data science workflow leveraging internet-sourced big data, ACS Catal. 13 (2023) 13267–13281, https://doi.org/10.1021/acscatal.3c01914.

[45] N. Dilshad, A. Ullah, J. Kim, et al., LocateUAV: unmanned aerial vehicle location estimation via contextual analysis in an iot Environment, IEEE Internet Things J. 10 (2023) 4021–4033, https://doi.org/10.1109/JIOT.2022.3162300.

[46] A. Mojiri, B.K. Walther, C. Jiang, et al., Telomerase therapy reverses vascular senescence and extends lifespan in progeria mice, Eur. Heart J. 42 (2021) 4352–4369, https://doi.org/10.1093/eurheartj/ehab547.

[47] N.S. Bobbitt, K.H. Shi, B.J. Bucior, et al., MOFX-DB: an online database of computational adsorption data for nanoporous materials, Chem. Eng. J. 68 (2023) 483–498, https://doi.org/10.1021/acs.jced.2c00583.

[48] R. Wang, H. Chen, S. Guo, et al., Machine learning toward realizing end-to-end biochar design for environmental remediation, ACS EST Eng. 4 (10) (2024) 2332–2345, https://doi.org/10.1021/acsestengg.4c00267.

[49] M.S. Khosrowshahi, H. Mashhadimoslem, H. Shayesteh, et al., Natural products derived porous carbons for CO_2 capture, Adv. Sci. 10 (2023) 2304289, https://doi.org/10.1002/advs.202304289.

[50] H. Wang, Y.Q. Li, X.Y. Xuan, et al., Machine learning accelerated discovery of covalent organic frameworks for environmental and energy applications, Environ. Sci. Technol. 59 (2025) 6361–6378, https://doi.org/10.1021/acs.est.5c00390.

[51] J. Plattfaut, M. Suckow, O. Klepel, et al., Modellierung und Simulation der templatgestützten Synthese von porösen Kohlenstoffgerüsten mittels COMSOL Multiphysics, Chem. Ing. Tech. 96 (2023) 318–328, https://doi.org/10.1002/cite.202300014.

[52] B.G. Fonseca, S.S. Thind, I. Booth, et al., A machine learning approach for the automated classification of bulk sp^2 and sp^3 carbon materials, J. Raman Spectrosc. 55 (2024) 15–25, https://doi.org/10.1002/jrs.6608.

[53] A. Edgar, Y.Y. Zhao, L. Palafox, et al., Porous metal properties analysis: a machine learning approach, JOM 73 (2021) 2039–2049, https://doi.org/10.1007/s11837-021-04695-x.

[54] H. Nguyen, T. Le, H. Nguyen, et al., Mechanical properties of AlCoCrCuFeNi high-entropy alloys using molecular dynamics and machine learning, Mater. Sci. Eng.: R: Rep. 160 (2024) 100833, https://doi.org/10.1016/j.mser.2024.100833.

[55] Y. Xu, Y.J. Yang, E. Wang, et al., Detect professional malicious user with metric learning in recommender systems, IEEE Trans. Knowl. Data Eng. 34 (2022) 4133–4146, https://doi.org/10.1109/TKDE.2020.3040618.

[56] Z. Yuan, H. Chen, P. Zhang, et al., A novel unsupervised approach to heterogeneous feature selection based on fuzzy mutual information, IEEE Trans. Fuzzy Syst. 30 (2022) 3395–3409, https://doi.org/10.1109/TFUZZ.2021.3114734.

[57] A. Jain, Machine learning in materials research: developments over the last decade and challenges for the future, Curr. Opin. Solid State Mater. Sci. 33 (2024) 101189, https://doi.org/10.1016/j.cossms.2024.101189.

[58] X. Liu, H.J. Peng, Toward next-generation heterogeneous catalysts: empowering surface reactivity prediction with machine learning, Engineering 39 (2024) 25–44, https://doi.org/10.1016/j.eng.2023.07.021.

[59] S.S. Jeyabalan, O.S. Ekande, B. Mainali, et al., A systematic review of recent advances in piezocatalysis-Synergetic heterojunctions for organic pollutants removal, immobilization, and scope of machine learning techniques, Chem. Eng. J. 498 (2024) 155086, https://doi.org/10.1016/j.cej.2024.155086.

[60] J. Zhang, C. Li, M.M. Rahaman, et al., A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches, Artif. Intell. Rev. 55 (2022) 2875–2944, https://doi.org/10.1007/s10462-021-10082-4.

[61] J.J. Yang, K. Hou, H. Wu, et al., Analysis of discharge image and particle characteristics of needle plate in transformer oil under AC voltage, IEEE Trans. Dielectr. Electr. Insul. 29 (2022) 1143–1151, https://doi.org/10.1109/TDEI.2022.3171183.

[62] W. Li, Y. Wang, Y. Shao, et al., TrackPuzzle: efficient registration of unlabeled PDR trajectories for learning indoor route graph, Future Gener. Comput. Syst. 149 (2023) 171–183, https://doi.org/10.1016/j.future.2023.07.019.

[63] K. Niemeijer, M. Mestdagh, P. Kuppens, et al., Tracking subjective sleep quality and mood with mobile sensing: multiverse study, J. Med. Internet Res. 24 (3) (2022) e25643, https://doi.org/10.2196/25643.

[64] L. Zhang, S.F. Shao, Image-based machine learning for materials science, J. Appl. Phys. 132 (2022) 100701, https://doi.org/10.1063/5.0087381.

[65] H. Chan, M. Cherukara, T.D. Loeffler, et al., Machine learning enabled autonomous microstructural characterization in 3D samples, npj Comput. Mater. 6 (1) (2020), https://doi.org/10.1038/s41524-019-0267-z.

[66] B.N. Han, Y.X. Lin, Y.F. Yang, et al., Deep-learning-enabled fast optical identification and characterization of 2D materials, Adv. Mater. 32 (2020) 2000953, https://doi.org/10.1002/adma.202000953.

[67] M. Shafiq, Z.Q. Gu, Deep residual learning for image recognition: a survey, Appl. Sci. 12 (2022) 8972, https://doi.org/10.3390/app12188972.

[68] K.X. Chen, G.X. Xu, J.Y. Qian, et al., A bypass-based U-Net for medical image segmentation, Intell. Sci. big Data Eng. Vis. Data Eng. 11935 (2019) 155–164, https://doi.org/10.1007/978-3-030-36189-1_13.

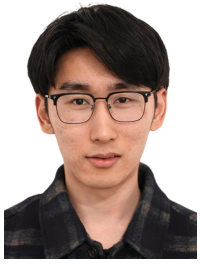
[69] M. Wasih, S. Ahmad, M. Almekkawy, A robust cascaded deep neural network for image reconstruction of single plane wave ultrasound RF data, Ultrasonics 132 (2023) 106981, https://doi.org/10.1016/j.ultras.2023.106981.
[70] N.S. Ross, P.M. Mashinini, C.S. Shibi, et al., A new intelligent approach of surface roughness measurement in sustainable machining of AM-316L stainless steel with deep learning models, Measurement 230 (2024) 114515, https://doi.org/10.1016/j.measurement.2024.114515.
[71] S. Agarwal, A. Sawant, M. Faisal, et al., Application of a deep learning semantic segmentation model to helium bubbles and voids in nuclear materials, Eng. Appl. Artif. Intell. 126 (2023) 106747, https://doi.org/10.1016/j.engappai.2023.106747.
[72] M. Saied, M. Raafat, S. Yehia, et al., Efficient pulmonary nodules classification using radiomics and different artificial intelligence strategies, Insights Imaging 14 (2023) 91, https://doi.org/10.1186/s13244-023-01441-6.
[73] T. Zhang, Y. Yang, A. Zhang, 3D reconstruction of porous media using a batch normalized variational auto-encoder, Comput. Geosci. 26 (2022) 1261–1278, https://doi.org/10.1007/s10596-022-10159-1.
[74] W.H. Xia, Y.L. Zhang, Y.J. Yang, et al., GAN inversion: a survey, IEEE Trans. Pattern Anal. Mach. Intell. 45 (2023) 3121–3138, https://doi.org/10.1109/TPAMI.2022.3181070.
[75] G.J. Ferdous, K.A. Sathi, M.A. Hossain, et al., SPT-swin: a shifted patch tokenization swin transformer for image classification, IEEE Access 12 (2024) 117617–117626, https://doi.org/10.1109/ACCESS.2024.3448304.
[76] A. Khan, Z. Rauf, A. Sohail, et al., A survey of the vision transformers and their CNN-transformer based variants, Artif. Intell. Rev. 56 (2023) 2917–2970, https://doi.org/10.1007/s10462-023-10595-0.
[77] K. Han, V.S. Sheng, Y.Q. Song, et al., Deep semi-supervised learning for medical image segmentation: a review, Expert Syst. Appl. 245 (2024) 123052, https://doi.org/10.1016/j.eswa.2023.123052.
[78] R. Espinosa, H. Ponce, O. Josue, A 3D orthogonal vision-based band-gap prediction using deep learning: a proof of concept, Comput. Mater. Sci. 202 (2022) 110967, https://doi.org/10.1016/j.commatsci.2021.110967.
[79] R. Alexander, A. Kaushal, P.T. Rao, et al., Identification and classification of disordered carbon materials in a composite matrix through machine learning approach integrated with Raman mapping, Diam. Relat. Mater. 142 (2024) 110741, https://doi.org/10.1016/j.diamond.2023.110741.
[80] B.M. Abraham, M.V. Jyothirmai, P. Sinha, et al., Catalysis in the digital age: unlocking the power of data with machine learning, Wiley Interdiscip. Rev. -Comput. Mol. Sci. 14 (2024) e1730, https://doi.org/10.1002/wcms.1730.
[81] H.X. Mai, T.C. Le, D. Chen, et al., Machine learning in the development of adsorbents for clean energy application and greenhouse gas capture, Adv. Sci. 9 (2022) 2203899, https://doi.org/10.1002/advs.202203899.
[82] C. Zuccarini, K. Ramachandran, D.D. Jayaseelan, Material discovery and modeling acceleration via machine learning, APL Mater. 12 (2024) 090601, https://doi.org/10.1063/5.0230677.
[83] D.A. Muyassiroh, F.A. Permatasari, F. Lskandar, Machine learning-driven advanced development of carbon-based luminescent nanomaterials, J. Mater. Chem. 10 (2022) 17431–17450, https://doi.org/10.1039/d2tc03789k.
[84] L. Zeng, Q. Lie, S. Shen, et al., Improved double deep Q network-based task scheduling algorithm in edge computing for makespan optimization, J. Mag. 29 (2024) 806–817, https://doi.org/10.26599/TST.2023.9010058.
[85] N. Ramasamy, A.J.L.P. Raj, V.V. Akula, et al., Leveraging experimental and computational tools for advancing carbon capture adsorbents research, Environ. Sci. Pollut. Res. 31 (2024) 55069–55098, https://doi.org/10.1007/s11356-024-34838-x.
[86] A. Merchant, S. Batzner, S.S. Schoenholz, et al., Scaling deep learning for materials discovery, Nature 624 (2023) 80–85, https://doi.org/10.1038/s41586-023-06735-9.
[87] P.C. Xu, X.B. Ji, M. Li, et al., Small data machine learning in materials science, NPJ Comput. Mater. 9 (2023) 42, https://doi.org/10.1038/s41524-023-01000-z.
[88] T. Toyao, A. Notes. Toward accelerated discovery of solid catalysts using extrapolative machine learning approach, Chem. Lett. 53 (2024) 163, https://doi.org/10.1093/chemle/upae163.
[89] H. Zhang, Y. Zhu, Q. Liu, et al., Preparation of porous carbon materials from biomass pyrolysis vapors for hydrogen storage, Appl. Energy 306 (2022) 118131, https://doi.org/10.1016/j.apenergy.2021.118131.
[90] A.H. Helaleh, M. Alizadeh, Performance prediction model of Miscible Surfactant-CO2 displacement in porous media using support vector machine regression with parameters selected by Ant colony optimization, J. Nat. Gas. Sci. Eng. 30 (2016) 388–404, https://doi.org/10.1016/j.jngse.2016.02.019.
[91] H. Mashhadimoslem, M.A. Bdol, P. Karimi, et al., Computational and machine learning methods for CO_2 capture using metal-organic frameworks, ACS Nano 18 (35) (2024) 23842–23875. ⟨https://doi.org/10.1021/acsnano.3c13001⟩.
[92] M. Osenberg, A. Hilger, M. Neumann, et al., Classification of FIB/SEM-tomography images for highly porous multiphase materials using random forest classifiers, J. Power Sources 570 (2023) 233030, https://doi.org/10.1016/j.jpowsour.2023.233030.
[93] R. Dubey, V. Guruviah, A data-driven approach for evaluation of electrolyte informatics on electrochemical performance of carbon-based electrode materials, Ionics 28 (2022) 2169–2183, https://doi.org/10.1007/s11581-022-04480-z.
[94] C. Wang, Q.X. Luo, E. Holm, Hierarchically structured classification of carbon nanostructures from TEM Images by machine learning and computer vision, Microsc. Microanal. 27 (2021) 448–449, https://doi.org/10.1017/S1431927621002105.
[95] M. Khanzadeh, S. Chowdhury, M. Marufuzzaman, et al., Porosity prediction: supervised-learning of thermal history for direct laser deposition, J. Manuf. Syst. 47 (2018) 69–82, https://doi.org/10.1016/j.jmsy.2018.04.001.
[96] V. Kumar, R.S. Singh, M. Rambabu, et al., Deep learning for hyperspectral image classification: a survey, Comput. Sci. Rev. 53 (2024) 100658, https://doi.org/10.1016/j.cosrev.2024.100658.
[97] H. Park, Y.H. Kang, J. Kim, Enhancing structure-property relationships in porous materials through transfer learning and cross-material few-shot learning, ACS Appl. Mater. Interfaces 15 (2023) 56375–56385, https://doi.org/10.1021/acsami.3c10323.
[98] F. Cipcigan, J. Booth, R.N.B. Ferreira, et al., Discovery of novel reticular materials for carbon dioxide capture using GFlowNets, Digit. Discov. 3 (2024) 449–455, https://doi.org/10.1039/D4DD00020J.
[99] M. Pal, P. Makauskas, S. Malik, Upscaling porous media using neural networks: a deep learning approach to homogenization and averaging, Processes 11 (2) (2023) 601, https://doi.org/10.3390/pr11020601.
[100] D. Kumbhar, A. Palliyarayil, D. Reghu, et al., Rapid discrimination of porous bio-carbon derived from nitrogen rich biomass using Raman spectroscopy and artificial intelligence methods, Carbon 178 (2021) 792–802, https://doi.org/10.1016/j.carbon.2021.03.064.
[101] A.V. Singh, M. Varma, M. Rai, et al., Advancing predictive risk assessment of chemicals via integrating machine learning, computational modeling, and chemical/nano-quantitative structure-activity relationship approaches, Adv. Intell. Syst. 6 (2024) 2300366, https://doi.org/10.1002/aisy.202300366.
[102] Y. Liu, X.Y. Li, Y.L. Chen, et al., Machine learning-encoded multiscale modelling and Bayesian optimization framework to design programmable metamaterials, Acta Mech. Sin. 41 (2025) 124061, https://doi.org/10.1007/s10409-024-24061-x.
[103] P. Ghorai, U. Mondal, A. Hazra, et al., Luminescent metal organic frameworks (LMOFs) and allied composites for the unveiling of organic environmental contaminants (explosive NACs, PAHs and EDCs) sensing through 'Molecular Recognition': a chronicle of recent penetration and future modelling, Coord. Chem. Rev. 518 (2024) 216085, https://doi.org/10.1016/j.ccr.2024.216085.
[104] M.S. Khosrowshahi, A.A. Aghajari, M. Rahimi, et al., Recent progress on advanced solid adsorbents for CO_2 capture: from mechanism to machine learning, Mater. Today Sustain. 27 (2024) 100900, https://doi.org/10.1016/j.mtsust.2024.100900.
[105] S.P. Fotias, I. Ismail, V. Gaganis, Optimization of well placement in carbon capture and storage (CCS): Bayesian optimization framework under permutation invariance, Appl. Sci. 14 (8) (2024) 3528, https://doi.org/10.3390/app14083528.
[106] C. Cleeton, A.H. Farmahini, L. Sarkisov, Performance-based ranking of porous materials for PSA carbon capture under the uncertainty of experimental data, Chem. Eng. J. 437 (2022) 135395, https://doi.org/10.1016/j.cej.2022.135395.
[107] A. Akinpelu, M. Bhullar, Y. Yao, Discovery of novel materials through machine learning, J. Phys.: Condens. Matter 36 (2024) 453001, https://doi.org/10.1088/1361-648X/ad6bdb.
[108] N.F. Alkayem, L. Shen, A. Mayya, et al., Prediction of concrete and FRC properties at high temperature using machine and deep learning: a review of recent advances and future perspectives, J. Build. Eng. 83 (2024) 108369, https://doi.org/10.1016/j.jobe.2023.108369.
[109] M. Tarawneh, A.J. Rami, Carbon dioxide evaporation heat transfer coefficient prediction in porous media using Machine learning algorithms based on experimental data, Therm. Sci. Eng. Prog. 55 (2024) 102929, https://doi.org/10.1016/j.tsep.2024.102929.
[110] S.S. Chong, Y.S. Ng, H.Q. Wang, et al., Advances of machine learning in materials science: Ideas and techniques, Front. Phys. 19 (2024) 13501, https://doi.org/10.1007/s11467-023-1325-z.
[111] S.Q. Wang, P. Xia, F.Y. Gong, et al., Multi objective optimization of recycled aggregate concrete based on explainable machine learning, J. Clean. Prod. 445 (2024) 141045, https://doi.org/10.1016/j.jclepro.2024.141045.
[112] Q. Liu, M.F. Jiang, J. Wang, et al., The rapid determination of three toxic ginkgolic acids in the decolorized process of ginkgo ketone ester based on raman spectroscopy and ResNeXt50 deep neural network, Chemosensors 12 (2024) 6, https://doi.org/10.3390/chemosensors12010006.
[113] S. Lin, Y. Zhuang, K. Chen, et al., Osteoinductive biomaterials: Machine learning for prediction and interpretation, Acta Biomater. 187 (2024) 422–433, https://doi.org/10.1016/j.actbio.2024.08.017.
[114] Y.H. Chen, S. Zhang, Y.H. Wen, et al., Prediction of material stability of two-dimensional semiconductors: an interpretable machine learning perspective, APL Mater. 12 (2024) 091120, https://doi.org/10.1063/5.0219418.
[115] P. Thakkar, S. Khatri, D. Dobariya, et al., Advances in materials and machine learning techniques for energy storage devices: a comprehensive review, J. Energy Storage 81 (2024) 110452, https://doi.org/10.1016/j.est.2024.110452.
[116] L. Bellucci, F. Delfino, V. Tozzini, In silico design, building and gas adsorption of nano-porous graphene scaffolds, Nanotechnology 32 (2020) 045704, https://doi.org/10.1088/1361-6528/abbe57.
[117] K. Kaneko, C. Ishil, M. Ruike, et al., Origin of superhigh surface area and microcrystalline graphitic structures of activated carbons, Carbon 30 (1992) 1075–1088, https://doi.org/10.1016/0008-6223(92)90139-N.
[118] M. Delpisheh, B. Ebrahimpour, A. Fattahi, et al., Leveraging machine learning in porous media, J. Mater. Chem. A 12 (2024) 20717–20782, https://doi.org/10.1039/D4TA00251B.
[119] X.H. Li, Z.H. Huang, S.S. Shao, et al., Machine learning prediction of physical properties and nitrogen content of porous carbon from agricultural wastes: effects

of activation and doping process, Fuel 356 (2024) 129623, https://doi.org/10.1016/j.fuel.2023.129623.
[120] J. Li, L. Pan, Y. Huang, et al., Biochar design for antibiotics adsorption via a hybrid machine-learning-based optimization framework, Sep. Purif. Technol. 348 (2024) 127666, https://doi.org/10.1016/j.seppur.2024.127666.
[121] R. Zou, Z.B. Yang, J.H. Zhang, et al., Machine learning application for predicting key properties of activated carbon produced from lignocellulosic biomass waste with chemical activation, Bioresour. Technol. 399 (2024) 130624, https://doi.org/10.1016/j.biortech.2024.130624.
[122] H.L. Li, Z.J. Ai, L.H. Yang, et al., Machine learning assisted predicting and engineering specific surface area and total pore volume of biochar, Bioresour. Technol. 369 (2023) 128417, https://doi.org/10.1016/j.biortech.2022.128417.
[123] L. Leng, L. Yang, X. Lei, et al., Machine learning predicting and engineering the yield, N content, and specific surface area of biochar derived from pyrolysis of biomass, Biochar 4 (2022) 63, https://doi.org/10.1007/s42773-022-00183-w.
[124] V.V. Korolev, A. Mitrofanov, E.I. Marchenko, et al., Transferable and extensible machine learning-derived atomic charges for modeling hybrid nanoporous materials, Chem. Mater. 32 (2020) 7822–7831, https://doi.org/10.1021/acs.chemmater.0c02468.
[125] S. Zhu, F.X. Sun, K.Z. Jiang, et al., Decoding intrinsic features of carbons from their capacitive performance, J. Electroanal. Chem. 961 (2024) 118253, https://doi.org/10.1016/j.jelechem.2024.118253.
[126] M. Khanzadeh, S. Chowdhury, M. Marufuzzaman, et al., Porosity prediction: supervised-learning of thermal history for direct laser deposition, J. Manuf. Syst. 47 (2018) 69–82, https://doi.org/10.1016/j.jmsy.2018.04.001.
[127] P.Z. Moghadam, Y.G. Chung, R.Q. Snurr, Progress toward the computational discovery of new metal-organic framework adsorbents for energy applications, Nat. Energy 9 (2024) 121–133, https://doi.org/10.1038/s41560-023-01417-2.
[128] K. Sarkar, M.A. Talukder, Structurally realistic carbide-derived carbon model in annealing molecular dynamics methodology with analytic bond-order potential, Mater. Adv. 5 (2024) 5738, https://doi.org/10.1039/d4ma00171k.
[129] H. Patel, J. Kröner, M. Schwan, et al., Molecular description of mechanical structure−property relationships of nanostructured porous carbon, J. Phys. Chem. C. 128 (2024) 21245–21252, https://doi.org/10.1021/acs.jpcc.4c07159.
[130] Y. Wang, Z. Fan, P. Qian, et al., Structure and pore size distribution in nanoporous carbon, Chem. Mater. 34 (2022) 617–628, https://doi.org/10.1021/acs.chemmater.1c03279.
[131] T.K. Stenczel, E. Zakariya, G. Liepuoniute, et al. Machine-learned acceleration for molecular dynamics in CASTEP. 2023, 159: 044803. ⟨https://doi.org/10.1063/5.0155621⟩.
[132] R. Su, Z. Li, X. Ma, et al. Machine learning, experiments and molecular simulation demonstrate the adsorption mechanism of acetone on porous carbon at different pressures. 2023, 323: 124480. ⟨https://doi.org/10.1016/j.seppur.2023.124480⟩.
[133] Y.F. Zhang, J. Tian, G. Li, et al., Design principles for gradient porous carbon on aqueous zinc-ion hybrid capacitors: a combined molecular dynamic and machine learning study, ACS Appl. Mater. Interfaces 17 (2025) 3448–3456, https://doi.org/10.1021/acsami.4c19397.
[134] X.Z. Yuan, J. Wang, S. Deng, et al., Recent advancements in sustainable upcycling of solid waste into porous carbons for carbon dioxide capture, Renew. Sustain. Energy Rev. 162 (2022) 112413, https://doi.org/10.1016/j.rser.2022.112413.
[135] M.R. Ketabchi, S. Babamohammadi, W.G. Davies, et al., Latest advances and challenges in carbon capture using bio-based sorbents: a state-of-the-art review, Carbon Capture Sci. Technol. 6 (2023) 100087, https://doi.org/10.1016/j.ccst.2022.100087.
[136] J. Xiao, X.F. Yuan, T.C. Zhang, et al., Nitrogen-doped porous carbon for excellent CO_2 capture: a novel method for preparation and performance evaluation, Sep. Purif. Technol. 298 (2022) 121602, https://doi.org/10.1016/j.seppur.2022.121602.
[137] C.L. Huang, H.Y. Yu, G.Z. Chen, et al., “Reinforced concrete” design of robust mineralized cellulose composite with multilayered structure for efficient CO_2 capture and passive radiative cooling ability, Compos. Sci. Technol. 258 (2024) 110886, https://doi.org/10.1016/j.compscitech.2024.110886.
[138] M.A. Shaida, S.S. Ansari, R. Muhammad, Prediction of CO_2 uptake in bio-waste based porous carbons using model agnostic explainable artificial intelligence, Fuel 380 (2025) 133183, https://doi.org/10.1016/j.fuel.2024.133183.
[139] D. Baskaran, P. Saravanan, L. Nagarajan, et al., An overview of technologies for capturing, storing, and utilizing carbon dioxide: technology readiness, large-scale demonstration, and cost, Chem. Eng. J. 491 (2024) 151998, https://doi.org/10.1016/j.cej.2024.151998.
[140] P.H.R. Amaral, D.T. Alvaro, D. Dionisio, et al., A machine learning model for adsorption energies of chemical species applied to CO_2 electroreduction, J. Electrochem. Soc. 169 (2022) 116505, https://doi.org/10.1149/1945-7111/ac9f7a.
[141] H. Wu, X. Wang, X. Wang, et al., Based on machine learning model for prediction of CO_2 adsorption of synthetic zeolite in two-step solid waste treatment, Arab. J. Chem. 17 (2) (2024) 105507, https://doi.org/10.1016/j.arabjc.2023.105507.
[142] H. Mashhadimoslem, M.A. Abdol, K. Zanganeh, et al., Development of the CO_2 adsorption model on porous adsorbent materials using machine learning algorithms, ACS Appl. Energy Mater. 7 (19) (2024) 8596–8609, https://doi.org/10.1021/acsaem.4c01465.
[143] X.Z. Zhu, D.C.W. Tsang, L. Wang, et al., Machine learning exploration of the critical factors for CO_2 adsorption capacity on porous carbon materials at different pressures, J. Clean. Prod. 173 (2020) 122915, https://doi.org/10.1016/j.jclepro.2020.122915.
[144] X.C. Ma, W.J. Xu, R.K. Su, et al., Insights into CO_2 capture in porous carbons from machine learning, experiments and molecular simulation, Sep. Purif. Technol. 306 (2023) 122521, https://doi.org/10.1016/j.seppur.2022.122521.
[145] A.G. Eslam, A. Ragab, T.M.Y. So, et al., Machine learning-assisted selection of adsorption-based carbon dioxide capture materials, J. Environ. Chem. Eng. 11 (2023) 110732, https://doi.org/10.1016/j.jece.2023.110732.
[146] J.J. Zhao, S.Y. Zhang, X.J. Zhang, et al., Machine learning and experimentally exploring the controversial role of nitrogen in CO_2 uptake by waste-derived nitrogen-containing porous carbons, Sci. Total Environ. 938 (2024) 173471, https://doi.org/10.1016/j.scitotenv.2024.173471.
[147] Z. Zhang, J.A. Schott, M. Liu, et al., Prediction of carbon dioxide adsorption via deep learning, Angew. Chem. Int. Ed. 58 (2019) 259–263, https://doi.org/10.1002/anie.201812363.
[148] P. Mehrmohammadi, A. Ghaemi, Investigating the effect of textural properties on CO_2 adsorption in porous carbons via deep neural networks using various training algorithms, Sci. Rep. 13 (2023) 21264, https://doi.org/10.1038/s41598-023-48683-4.
[149] K. Palle, S. Vunguturi, S.N. Gayatri, et al., The prediction of -CO_2 adsorption on rice husk activated carbons via deep learning neural network, Communications 12 (2022) 434–440, https://doi.org/10.1557/s43579-022-00197-2.
[150] M. Maheri, C. Bazan, S. Zendehboudi, et al., Machine learning to assess CO_2 adsorption by biomass waste, J. CO_2 Util. 76 (2023) 102590, https://doi.org/10.1016/j.jcou.2023.102590.
[151] J. Cui, F. Wu, W. Zhang, et al., Direct prediction of gas adsorption via spatial atom interaction learning, Nat. Commun. 14 (2023) 7043, https://doi.org/10.1038/s41467-023-42863-6.
[152] C.B. Peres, L.C.D. Morais, P.M.R. Resende, Carbon adsorption on waste biomass of passion fruit peel: a promising machine learning model for CO_2 capture, J. CO_2 Util. 80 (2024) 102680, https://doi.org/10.1016/j.jcou.2024.102680.
[153] K. Huo, Y. Sun, H. Jiang, et al., Petroleum pitch-derived porous carbon materials as metal-free catalyst for dry reforming of methane, Molecules 29 (2024) 4642, https://doi.org/10.3390/molecules29194642.
[154] E.H. Cho, L.C. Lin, Nanoporous material recognition via 3D convolutional neural networks: Prediction of adsorption properties, J. Phys. Chem. 12 (2021) 2279–2285, https://doi.org/10.1021/acs.jpclett.1c00293.
[155] C. Zhang, D. Li, Y.C. Xie, et al., Machine learning assisted rediscovery of methane storage and separation in porous carbon from material literature, Fuel 290 (2021) 120080, https://doi.org/10.1016/j.fuel.2020.120080.
[156] H. Wang, M. Zhang, X. Xia, et al., Lattice Boltzmann prediction of CO_2 and CH_4 competitive adsorption in shale porous media accelerated by machine learning for CO_2 sequestration and enhanced CH_4 recovery, Appl. Energy 370 (2024) 12368, https://doi.org/10.1016/j.apenergy.2024.123638.
[157] M. Rahimi, A. Mehrpanah, P. Mouchani, et al., Optimizing methane uptake on N/O functionalized graphene via DFT, machine learning, and uniform manifold approximation and projection (UMAP) techniques, Ind. Eng. Chem. Res. 63 (44) (2024) 18940–18956, https://doi.org/10.1021/acs.iecr.4c02626.
[158] M. Tavakolian, N.S. Reza, N. Chen, et al., Modeling of methane and carbon dioxide sorption capacity in tight reservoirs using Machine learning techniques, Fuel 360 (2024) 130578, https://doi.org/10.1016/j.fuel.2023.130578.
[159] X. Zhang, D. Ren, X.L. Fu, et al., A novel dual-way inference modeling method for coal coking: predicting H_2 and CH_4 concentrations in coke oven gas and inferring optimal reaction conditions, Fuel 381 (2025) 133325, https://doi.org/10.1016/j.fuel.2024.133325.
[160] M.S.M. Zaini, M. Arshad, A.S. Syed, Adsorption isotherm and kinetic study of methane on palm kernel shell-derived activated carbon, J. Bioresour. Bioprod. 8 (2023) 66–77, https://doi.org/10.1016/j.jobab.2022.11.002.
[161] C.L. Cui, J. Tang, H. Xia, et al., Virtual sample generation method based on generative adversarial fuzzy neural network, Neural Comput. Appl. 35 (2023) 6979–7001, https://doi.org/10.1007/s00521-022-08104-5.
[162] K.S. Chandana, R. Kamesh, Machine learning insights into prediction of H_2 gravimetric capacity in Mg-based pure metal alloys, Int. J. Hydrog. Energy 77 (2024) 695–711, https://doi.org/10.1016/j.ijhydene.2024.06.186.
[163] G. Klepp, Modelling activated carbon hydrogen storage tanks using machine learning models, Energy 306 (2024) 132318, https://doi.org/10.1016/j.energy.2024.132318.
[164] L. Zhang, M.D. Allendorf, B.X. Rafael, et al., Fundamentals of hydrogen storage in nanoporous materials, Prog. Energy 4 (2022) 042013, https://doi.org/10.1088/2516-1083/ac8d44.
[165] Y.H. Xu, Y. Zhou, C.Q. Li, et al., Unraveling the potential of solid-state hydrogen storage materials: insights from first principle calculations, Fuel 373 (2024) 132340, https://doi.org/10.1016/j.fuel.2024.132340.
[166] A.I. Osman, M. Nasr, A.S. Eltaweil, et al., Advances in hydrogen storage materials: harnessing innovative technology, from machine learning to computational chemistry, for energy storage solutions, Int. J. Hydrog. Energy 67 (2024) 1270–1294, https://doi.org/10.1016/j.ijhydene.2024.03.223.
[167] M.I.M. Kusdhang, S.M. Lyth, New insights into hydrogen uptake on porous carbon materials via explainable machine learning, Carbon 179 (2021) 190–201, https://doi.org/10.1016/j.carbon.2021.04.036.
[168] H.V. Thanh, S.E. Taremsari, B. Ranjbar, et al., Hydrogen storage on porous carbon adsorbents: rediscovery by nature-derived algorithms in random forest machine learning model, Energies 16 (5) (2023) 2348, https://doi.org/10.3390/en16052348.
[169] A. Gotzias, E. Tylianakis, G. Froudakis, et al., Theoretical study of hydrogen adsorption in oxygen functionalized carbon slit pores, Microporous Mesoporous Mater. 154 (2012) 38–44, https://doi.org/10.1016/j.micromeso.2011.10.011.

[170] S. Schaefer, A. Jeder, G. Sdanghi, et al., Oxygen-promoted hydrogen adsorption on activated and hybrid carbon materials, Microporous Mesoporous Mater. 154 (2012) 38–44, https://doi.org/10.1016/j.micromeso.2011.10.011.
[171] G. Sethia, A. Sayari, Activated carbon with optimum pore size distribution for hydrogen storage, Carbon 99 (2016) 289–294, https://doi.org/10.1016/j.carbon.2015.12.032.
[172] Z. Jia, S. Lu, P. Song, et al., Machine learning accelerates design of bilayer-modified graphene hydrogen storage materials, Sep. Purif. Technol. 352 (2025) 128229, https://doi.org/10.1016/j.seppur.2024.128229.
[173] H. Hadiyawarman, D. Aji, G.E. Timuda, et al., Bibliometric review of carbonaceous-based sorbent for hydrogen storage, Clean. Techn Environ. Policy (2024), https://doi.org/10.1007/s10098-024-03096-3.
[174] D. Xu, J. Yang, Z. Xu, et al., Accelerated discovery of CO_2 solid sorbents using active machine learning: review and perspectives, Energy Fuels 38 (2024) 17161–17178, https://doi.org/10.1021/acs.energyfuels.4c02789.
[175] G. Zheng, S. Zhang, L. Meng, et al., Machine learning-guided design and synthesis of eco-friendly poly(ethylene oxide) membranes for high-efficacy CO_2/N_2 separation, Adv. Funct. Mater. 34 (2024) 2410075, https://doi.org/10.1002/adfm.202410075.
[176] S. Wang, Z. Zhang, S. Dai, et al., Insights into CO_2/N_2 selectivity in porous carbons from deep learning, ACS Mater. Lett. 1 (2019) 558–563, https://doi.org/10.1021/acsmaterialslett.9b00374.
[177] S. Wang, Y. Li, S. Dai, et al., Prediction by convolutional neural networks of CO_2/N_2 selectivity in porous carbons from N_2 adsorption isotherm at 77K, Angew. Chem. Int. Ed. 59 (2020) 19645–19648, https://doi.org/10.1002/anie.202005931.
[178] B. Li, S. Wang, Z.Q. Tian, et al., Understanding the $CO_2/CH_4/N_2$ separation performance of nanoporous amorphous N-doped carbon combined hybrid Monte Carlo with machine learning, Adv. Theory Simul. 5 (2021) 2100378, https://doi.org/10.1002/adts.202100378.
[179] Y. Pan, L. He, Y. Ren, et al., Analysis of influencing factors on the gas separation performance of carbon molecular sieve membrane using machine learning technique, Membranes 12 (2022) 100, https://doi.org/10.3390/membranes12010100.
[180] K. Hippalgaonkar, Q. Li, X. Wang, et al., Knowledge-integrated machine learning for materials: lessons from gameplaying and robotics, Nat. Rev. Mater. 8 (2023) 241–260, https://doi.org/10.1038/s41578-022-00513-1.
[181] A. Chen, Z. Wang, K.L.L. Vidaurre, et al., Knowledge-reused transfer learning for molecular and materials science, J. Energy Chem. 98 (2024) 149–168, https://doi.org/10.1016/j.jechem.2024.06.013.
[182] Q.A. Sial, U. Safder, S. Iqbal, et al., Advancement in supercapacitors for IoT applications by using machine learning: Current trends and future technology, Sustainability 16 (2024) 1516, https://doi.org/10.3390/su16041516.
[183] S.A.A. Shah, R. Idrees, S. Saeed, A critical review on polyimide derived carbon materials for high-performance supercapacitor electrodes, J. Energy Storage 55 (2022) 105667, https://doi.org/10.1016/j.est.2022.105667.
[184] H. Su, S. Lin, S. Deng, et al., Predicting the capacitance of carbon-based electric double layer capacitors by machine learning, Nanoscale Adv. 1 (2019) 2162, https://doi.org/10.1039/c9na00105k.
[185] S. Mishra, R. Srivastava, A. Muhammad, et al., The impact of physicochemical features of carbon electrodes on the capacitive performance of supercapacitors: a machine learning approach, Sci. Rep. 13 (2023) 6494, https://doi.org/10.1038/s41598-023-33524-1.
[186] X.R. Liu, H.P. Yang, P.X. Xue, et al., Machine learning modeling of the capacitive performance of N-doped porous biochar electrodes with experimental verification, Renew. Energy 231 (2024) 120969, https://doi.org/10.1016/j.renene.2024.120969.
[187] W.A.M. Wickramaaarachchi, M. Minakshi, X. Gao, et al., Hierarchical porous carbon from mango seed husk for electro-chemical energy storage, Chem. Eng. J. Adv. 8 (2021) 100158, https://doi.org/10.1016/j.ceja.2021.100158.
[188] H. Liu, Z. Cui, Z. Qiao, et al., Machine learning-assisted prediction, screen, and interpretation of porous carbon materials for high-performance supercapacitors, J. Mater. Inform. 4 (2024) 16, https://doi.org/10.20517/jmi.2024.29.
[189] X. Yang, C. Yuan, S. He, et al., Machine learning prediction of specific capacitance in biomass derived carbon materials: effects of activation and biochar characteristics, Fuel 331 (2023) 125718, https://doi.org/10.1016/j.fuel.2022.125718.
[190] P. Liu, Y.P. Wen, L. Huang, et al., An emerging machine learning strategy for the assisted-design of high-performance supercapacitor materials by mining the relationship between capacitance and structural features of porous carbon, J. Electroanal. Chem. 899 (2021) 115684, https://doi.org/10.1016/j.jelechem.2021.115684.
[191] W.Z. Tawfik, S.N. Mohammad, K.H. Rahouma, et al., An artificial neural network model for capacitance prediction of porous carbon-based supercapacitor electrodes, J. Energy Storage 73 (2023) 108830, https://doi.org/10.1016/j.est.2023.108830.
[192] W.Z. Tawfik, S.N. Mohammad, K.H. Rahouma, et al., Machine learning models for capacitance prediction of porous carbon-based supercapacitor electrodes, Phys. Scr. 99 (2024) 026001, https://doi.org/10.1088/1402-4896/ad190c.
[193] T. Wang, R. Pan, M.L. Maryins, et al., Machine-learning-assisted material discovery of oxygen-rich highly porous carbon active materials for aqueous supercapacitors, Nat. Commun. 14 (2023) 4607, https://doi.org/10.1038/s41467-023-40282-1.
[194] D. Golze, M. Hirvensalo, H.L. Patircia, et al., Accurate computational prediction of core-electron binding energies in carbon-based materials: a machine-learning model combining density-functional theory and GW, Chem. Mater. 34 (2022) 6240–6254, https://doi.org/10.1021/acs.chemmater.1c04279.
[195] C. Zhao, X. Lu, H. Tu, et al., Research on specific capacitance prediction of biomass carbon-based supercapacitors based on machine learning, J. Energy Storage 97 (2024) 112974, https://doi.org/10.1016/j.est.2024.112974.
[196] X.R. Liu, H.P. Yang, Y.J. Tang, et al., Deep learning prediction and experimental investigation of specific capacitance of nitrogen-doped porous biochar, Bioresour. Technol. 403 (2024) 130865, https://doi.org/10.1016/j.biortech.2024.130865.
[197] D. Li, A. Liang, M. Zhou, et al., Energy utilization of agricultural waste: machine learning prediction and pyrolysis transformation, Waste Manag. 175 (2024) 235–244, https://doi.org/10.1016/j.wasman.2024.01.003.
[198] M. Akdemir, D.E. Karakas, M. Kaya, Synthesis of a dual-functionalized carbon-based material as catalyst and supercapacitor for efficient hydrogen production and energy storage: Pd-supported pomegranate peel, Energy Storage 4 (2022) e284, https://doi.org/10.1002/est2.284.
[199] T. Weigert, Q. Tian, K. Lian, State-of-charge prediction of batteries and battery–supercapacitor hybrids using artificial neural networks, J. Power Sources 196 (2011) 4061–4066, https://doi.org/10.1016/j.jpowsour.2010.10.075.
[200] V. Sawant, R. Deshmukh, C. Awati, Machine learning techniques for prediction of capacitance and remaining useful life of supercapacitors: a comprehensive review, J. Energy Chem. 77 (2023) 438–451, https://doi.org/10.1016/j.jechem.2022.11.012.
[201] Y. Zhou, Y. Huang, J. Pang, et al., Remaining useful life prediction for supercapacitor based on long short-term memory neural network, J. Power Sources 440 (2019) 227149, https://doi.org/10.1016/j.jpowsour.2019.227149.
[202] C. Liu, Y. Zhang, J. Sun, et al., Stacked bidirectional LSTM RNN to evaluate the remaining useful life of supercapacitor, Int. J. Energy Res. 46 (2022) 3034–3043, https://doi.org/10.1002/er.7360.
[203] C.L. Liu, D. Li, L. Wang, et al., Strong robustness and high accuracy in predicting remaining useful life of supercapacitors, APL Mater. 10 (2022) 061106, https://doi.org/10.1063/5.0092074.
[204] C.X. Wang, R. Xiong, J.P. Tian, et al., Rapid ultracapacitor life prediction with a convolutional neural network, Appl. Energy 305 (2022) 117819, https://doi.org/10.1016/j.apenergy.2021.117819.
[205] Y.X. Sun, P.H. Sun, J.X. Jia, et al., Machine learning in clarifying complex relationships: biochar preparation procedures and capacitance characteristics, Chem. Eng. J. 485 (2024) 149975, https://doi.org/10.1016/j.cej.2024.149975.
[206] D.V. Krasnikov, E.M. Khabushev, A. Gaev, et al., Machine learning methods for aerosol synthesis of single-walled carbon nanotubes, Carbon 202 (2023) 76–82, https://doi.org/10.1016/j.carbon.2022.10.044.
[207] F. Guo, H. Lv, X. Wu, et al., A machine learning method for prediction of remaining useful life of supercapacitors with multi-stage modification, J. Energy Storage 73 (2023) 109160, https://doi.org/10.1016/j.est.2023.109160.
[208] M. Jiao, D.Q. Wang, J.L. Qiu, A GRU-RNN based momentum optimized algorithm for SOC estimation, J. Power Sources 459 (2020) 228051, https://doi.org/10.1016/j.jpowsour.2020.228051.
[209] W. Zhang, Q. Chen, J.F. Chen, et al., Machine learning for hydrothermal treatment of biomass: a review, Bioresour. Technol. 370 (2023) 128547, https://doi.org/10.1016/j.biortech.2022.128547.
[210] J.J. Liu, Y.W. You, L. Huang, et al., Precisely tunable instantaneous carbon rearrangement enables low-working-potential hard carbon toward sodium-ion batteries with enhanced energy density, Adv. Mater. 36 (2024) 2407369, https://doi.org/10.1002/adma.202407369.
[211] N. Zhang, B. Yang, K. Liu, et al., Machine learning in screening high performance electrocatalysts for CO_2 reduction, Small Methods 5 (2021) 2100987, https://doi.org/10.1002/smtd.202100987.
[212] Y.Z. Wang, Z.Y. Fan, P. Qian, et al., Structure and pore size distribution in nanoporous carbon, Chem. Mater. 34 (2022) 617–628, https://doi.org/10.1021/acs.chemmater.1c03279.
[213] F.Z. Azar, A.E. Kasmi, L.Á. Maria, et al., Selective biomass conversion over novel designed tandem catalyst, J. Bioresour. Bioprod. 9 (2024) 508–517, https://doi.org/10.1016/j.jobab.2024.09.001.
[214] K. Pobłocki, M. Pawlak, J. Drzeżdżon, et al., Catalytic materials based on metals (ions) used in the upcycling of plastics and polymers into fuels and valuable chemicals as part of sustainable development, Mater. Sci. Eng.: R: Rep. 162 (2025) 100881, https://doi.org/10.1016/j.mser.2024.100881.
[215] X. Yang, H. He, J. Qiu, Fabrication of biomass-based functional carbon materials for energy conversion and storage, Mater. Sci. Eng.: R: Rep. 154 (2023) 100736, https://doi.org/10.1016/j.mser.2023.100736.
[216] X.Z. Zhu, Z.H. Wan, D.C.W. Tsang, et al., Machine learning for the selection of carbon-based materials for tetracycline and sulfamethoxazole adsorption, Chem. Eng. J. 406 (2021) 126782, https://doi.org/10.1016/j.cej.2020.126782.
[217] F.V. Christian, L.C. Moreno, Y.A. Yazmin, Adsorption of pharmaceutical pollutants on $ZnCl_2$-activated biochar from corn cob: efficiency, selectivity and mechanism, J. Bioresour. Bioprod. 9 (2024) 58–73, https://doi.org/10.1016/j.jobab.2023.10.003.
[218] K.H.H. Aziz, F.S. Mustafa, M.A. Hassan, et al., Biochar as green adsorbents for pharmaceutical pollution in aquatic environments: a review, Desalination 583 (2024) 117725, https://doi.org/10.1016/j.desal.2024.117725.
[219] H.S. Shahraki, R. Bushra, N. Shakeel, et al., Papaya peel waste carbon dots/reduced graphene oxide nanocomposite: from photocatalytic decomposition of methylene blue to antimicrobial activity, J. Bioresour. Bioprod. 8 (2023) 162–175, https://doi.org/10.1016/j.jobab.2023.01.009.
[220] R. Kushwaha, M.K. Singh, S. Krishnan, et al., Machine learning enabled property prediction of carbon-based electrodes for supercapacitors, J. Mater. Sci. 58 (2023) 15448–15458, https://doi.org/10.1007/s10853-023-08981-8.

[221] R. Dubey, V. Guruviah, Machine learning enabled performance prediction of biomass-derived electrodes for asymmetric supercapacitor, Futur. Commun. Netw. Technol. 995 (2021), https://doi.org/10.1007/978-981-19-9748-8_42.
[222] R. Kushwaha, M.K. Singh, S. Krishnan, et al., Machine learning enabled property prediction of carbon-based electrodes for supercapacitors, J. Mater. Sci. 58 (2023) 15448–15458, https://doi.org/10.1007/s10853-023-08981-8.
[223] Y. Ma, J. Gao, X. Chen, et al., Regulation of the mesopore proportion of porous carbon for optimizing the performance of electric double layer capacitors, J. Energy Storage 35 (2021) 102299, https://doi.org/10.1016/j.est.2021.102299.
[224] D. Liu, T. Sun, Y. Hu, et al., Precisely tailoring pore structure in sunflower plate-derived N, O co-doped carbons for high-performance supercapacitors, J. Ind. Eng. Chem. 125 (2023) 284–293, https://doi.org/10.1016/j.jiec.2023.05.038.
[225] P. Qi, J.X. Wang, L. Yang, et al., Self-assembled nano Co-Ni pompon structured hierarchical porous carbon and the excellent electrochemical capacity, J. Alloy. Compd. 960 (2023) 170804, https://doi.org/10.1016/j.jallcom.2023.170804.
[226] A. Ghosh, Y.H. Lee, Carbon-based electrochemical capacitors, ChemSusChem 5 (2012) 480–499, https://doi.org/10.1002/cssc.201100645.
[227] M. Rahimi, M.H. Abbaspour, A. Rohani, et al., Synergetic effect of N/O functional groups and microstructures of activated carbon on supercapacitor performance by machine learning, J. Power Sources 521 (2022) 230968, https://doi.org/10.1016/j.jpowsour.2021.230968.
[228] M. Zhou, A. Gallegos, K. Liu, et al., Insights from machine learning of carbon electrodes for electric double layer capacitors, Carbon 157 (2020) 147–152, https://doi.org/10.1016/j.carbon.2019.08.090.
[229] Z. Supiyeva, X. Pan, Q. Abbas, The critical role of nanostructured carbon pores in supercapacitors, Curr. Opin. Electrochem. 39 (2023) 101249, https://doi.org/10.1016/j.coelec.2023.101249.
[230] P. Kowalczyk, A.P. Terzyk, P. Erwardt, et al., Machine learning-assisted design of porous carbons for removing paracetamol from aqueous solutions, Carbon 198 (2022) 371–381, https://doi.org/10.1016/j.carbon.2022.07.029.
[231] X.Z. Yuan, M. Suvarna, S. Low, et al., Applied machine learning for prediction of CO_2 adsorption on biomass waste-derived porous carbons, Environ. Sci. Technol. 55 (2021) 11925–11936, https://doi.org/10.1021/acs.est.1c01849.
[232] E.M. Chan, C. Xu, A.W. Mao, et al., Reproducible, high-throughput synthesis of colloidal nanocrystals for optimization in multidimensional parameter space, Nano Lett. 10 (2010) 1874–1885, https://doi.org/10.1021/nl100669s.
[233] N.J. Szymanski, B. Rendy, Y. Fei, et al., An autonomous laboratory for the accelerated synthesis of novel materials, Nature 624 (2023) 86–91, https://doi.org/10.1038/s41586-023-06734-w.
[234] T. Dai, S. Vijayakrishnan, F.T. Szczypiński, et al., Autonomous mobile robots for exploratory synthetic chemistry, Nature 635 (2024) 890–897, https://doi.org/10.1038/s41586-024-08173-7.
[235] T. Ha, D. Lee, Y. Kwon, et al., AI-driven robotic chemist for autonomous synthesis of organic molecules, Sci. Adv. 9 (44) (2023), https://doi.org/10.1126/sciadv.adj0461.
[236] H.C. Tao, T.Y. Wu, M. Aldeghi, et al., Nanoparticle synthesis assisted by machine learning, Nat. Rev. Mater. 6 (2021) 701–716, https://doi.org/10.1038/s41578-021-00337-5.
[237] P. Nikolaev, D. Hooper, F. Webber, et al., Autonomy in materials research: a case study in carbon nanotube growth, Comput. Mater. 2 (2016) 16031, https://doi.org/10.1038/npjcompumats.2016.31.
[238] R. Rao, C. Jennifer, P. Nikolaev, et al., Advanced machine learning decision policies for diameter control of carbon nanotubes, Comput. Mater. 7 (2021) 157, https://doi.org/10.1038/s41524-021-00629-y.
[239] M. Delpisheh, B. Ebrahimpour, A. Fattahi, et al., Leveraging machine learning in porous media, J. Mater. Chem. A 12 (2024) 20717–20782, https://doi.org/10.1039/D4TA07195F.
[240] H. Furukawa, N. Ko, Y.B. Go, et al., Ultrahigh porosity in metal-organic frameworks, Science 329 (5990) (2010), https://doi.org/10.1126/science.1192160.
[241] H. Furukawa, K.E. Cordova, M. O'Keeffe, et al., The chemistry and applications of metal-organic frameworks, Science 341 (6149) (2013), https://doi.org/10.1126/science.1230444.
[242] A.J. Howarth, Y. Liu, P. Li, et al., Chemical, thermal and mechanical stabilities of metal–organic frameworks, Nat. Rev. Mater. 1 (2016) 15018, https://doi.org/10.1038/natrevmats.2015.18.
[243] J.J. Low, A.I. Benin, P. Jakubczak, et al., Virtual high throughput screening confirmed experimentally: porous coordination polymer hydration, J. Am. Chem. Soc. 131 (2009) 15834–15842, https://doi.org/10.1021/ja9061344.
[244] S. Liu, H. Zhang, Q. Zhao, et al., Metal-organic framework derived nitrogen-doped porous carbon@graphene sandwich-like structured composites as bifunctional electrocatalysts for oxygen reduction and evolution reactions, Carbon 106 (2016) 74–83, https://doi.org/10.1016/j.carbon.2016.05.021.
[245] M.A. Akasheh, E.F. Malik, O. Hujran, et al., A decade of research on machine learning techniques for predicting employee turnover: A systematic literature review, Expert Syst. Appl. 238 (2024) 121794, https://doi.org/10.1016/j.eswa.2023.121794.
[246] Y. Huang, J. Li, M. Li, et al., Application of machine learning in predicting survival outcomes involving real-world data: a scoping review, BMC Med. Res. Methodol. 23 (2023) 268, https://doi.org/10.1186/s12874-023-02078-1.
[247] S. Demir, E.K. Sahin, Comparison of tree-based machine learning algorithms for predicting liquefaction potential using canonical correlation forest, rotation forest, and random forest based on CPT data, Soil Dyn. Earthq. Eng. 154 (2022) 107130, https://doi.org/10.1016/j.soildyn.2021.107130.
[248] I. Kanjanasurat, W. Jungsuwadee, A. Lasakul, et al., Comparison of logistic regression and random forest algorithms for airport's runway assignment, J. Phys.: Conf. Ser. 2497 (2023) 012016, https://doi.org/10.1088/1742-6596/2497/1/012016.
[249] M.A. Bayir, K. Shamsi, H. Kaynak, et al., Topological forest, IEEE Access 10 (2022) 131711–131721, https://doi.org/10.1109/ACCESS.2022.3229008.
[250] M.M. Forootan, I. Larki, R. Zahedi, et al., Machine learning and deep learning in energy systems: a review, Sustainability 14 (8) (2022) 4832, https://doi.org/10.3390/su14084832.
[251] J. Kufel, B. Katarzyna, S. Kocot, et al., What is machine learning, artificial neural networks and deep learning?-Examples of practical applications in medicine, Diagnostics 13 (15) (2023) 2582, https://doi.org/10.3390/diagnostics13152582.
[252] W. Samek, G. Montavon, S. Lapuschkin, C.J. Anders, et al., Explaining deep neural networks and beyond: a review of methods and applications, Proc. IEEE 109 (2021) 247–278, https://doi.org/10.1109/JPROC.2021.3060483.
[253] M. Pichler, F. Hartig, Machine learning and deep learning-A review for ecologists, Methods Ecol. Evol. 14 (2023) 994–1016, https://doi.org/10.1111/2041-210X.14061.
[254] D. Drikakis, F. Sofos, Can artificial intelligence accelerate fluid mechanics research? Fluids 8 (7) (2023) 212, https://doi.org/10.3390/fluids8070212.
[255] L. Alzubaidi, J. Zhang, A.J. Humaidi, et al., Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, J. Big Data 8 (2021) 53, https://doi.org/10.1186/s40537-021-00444-8.
[256] S. Uppalapati, P. Paramasivam, N. Kilari, et al., Precision biochar yield forecasting employing random forest and XGBoost with Taylor diagram visualization, Sci. Rep. 15 (2025) 7105, https://doi.org/10.1038/s41598-025-91450-w.
[257] N.S. Rajpurohit, P.K. Kamani, M. Lenka, et al., Predictive modeling of product yields in microwave-assisted co-pyrolysis of biomass and plastic with enhanced interpretability using explainable AI approaches, J. Anal. Appl. Pyrolysis 188 (2025) 107021, https://doi.org/10.1016/j.jaap.2025.107021.
[258] B. Dou, Z. Zhu, E. Merkurjev, et al., Machine learning methods for small data challenges in molecular science 123 (2023) 8736–8780, https://doi.org/10.1021/acs.chemrev.3c00189.
[259] X. Rao, S. Zhang, J.J. Zhang, Carbon semi-tubes for electrochemical energy catalysis, Electrochem. Energy Rev. 8 (2025) 7, https://doi.org/10.1007/s41918-025-00238-z.
[260] Z. Huang, K. Huang, J. Zhang, et al., Active temperature control strategy in adsorption-based carbon capture: principle, methodology and case study, Sep. Purif. Technol. 364 (2025) 132581, https://doi.org/10.1016/j.seppur.2025.132581.
[261] J. Li, B. Michalkiewicz, C. Ma, et al., Selective preparation of biomass-derived porous carbon with controllable pore sizes toward highly efficient CO_2 capture, Chem. Eng. J. 360 (2019) 250–259, https://doi.org/10.1016/j.cej.2018.11.204.
[262] X. Yu, S. Liu, G. Lin, et al., Promotion effect of KOH surface etching on sucrose-based hydrochar for acetone adsorption, Appl. Surf. Sci. 496 (2019) 143617, https://doi.org/10.1016/j.apsusc.2019.143617.
[263] S. Li, H. Zhang, S. Hu, et al., Synthesis of hierarchical porous carbon in molten salt and its application for dye adsorption, Nanomaterials 9 (8) (2019) 1098, https://doi.org/10.3390/nano9081098.
[264] S. Wang, H. Sun, H.M. Ang, et al., Adsorptive remediation of environmental pollutants using novel graphene-based nanomaterials, Chem. Eng. J. 226 (2013) 336–347, https://doi.org/10.1016/j.cej.2013.04.070.
[265] C. Xie, Y.C. Xie, C. Zhang, et al., Explainable machine learning for carbon dioxide adsorption on porous carbon, J. Environ. Chem. Eng. 11 (2023) 109053, https://doi.org/10.1016/j.jece.2022.109053.
[266] M. Karimi, M. Shirzad, J.A.C. Silva, et al., Carbon dioxide separation and capture by adsorption: a review, Environ. Chem. Lett. 21 (2023) 2041–2084, https://doi.org/10.1007/s10311-023-01589-z.
[267] Z. Chen, K.O. Kirlikovali, K.B. Idrees, et al., Porous materials for hydrogen storage, Chem 8 (2022) 693–716, https://doi.org/10.1016/j.chempr.2022.01.012.
[268] U. Kamran, S. Park, Chemically modified carbonaceous adsorbents for enhanced CO_2 capture: a review, J. Clean. Prod. 290 (2021) 125776, https://doi.org/10.1016/j.jclepro.2020.125776.
[269] Y. Zhang, L. Liu, P. Zhang, et al., Ultra-high surface area and nitrogen-rich porous carbons prepared by a low-temperature activation method with superior gas selective adsorption and outstanding supercapacitance performance, Chem. Eng. J. 355 (2019) 109–319, https://doi.org/10.1016/j.cej.2018.08.169.
[270] A.L. Osman, A. Walaa, M. Nasr, et al., Enhanced hydrogen storage efficiency with sorbents and machine learning: a review, Environ. Chem. Lett. 22 (2024) 1703–1740, https://doi.org/10.1007/s10311-024-01741-3.

[271] G. Singh, K. Ramadass, V.D.B.C. Dasireddy, et al., Material-based generation, storage, and utilisation of hydrogen, Prog. Mater. Sci. 135 (2023) 101104, https://doi.org/10.1016/j.pmatsci.2023.101104.
[272] A. Rehman, G. Nazir, K.Y. Rhee, et al., A rational design of cellulose-based heteroatom-doped porous carbons: promising contenders for CO_2 adsorption and separation, Chem. Eng. J. 420 (2021) 130421, https://doi.org/10.1016/j.cej.2021.130421.
[273] M.S.B. Reddy, D. Ponnamma, K.K. Sadasivuni, et al., Carbon dioxide adsorption based on porous materials, RCS Adv. 11 (2021) 12658–12681, https://doi.org/10.1039/D0RA10902A.
[274] A.H. Farmahini, S. Krishnamurthy, D. Friedrich, et al., Performance-based screening of porous materials for carbon capture, Chem. Rev. 121 (2021) 10666–10741, https://doi.org/10.1021/acs.chemrev.0c01266.
[275] M.Y. Kurbakov, V.V. Sulimova, A.V. Kopylov, et al., Determining the orderliness of carbon materials with nanoparticle imaging and explainable machine learning, Nanoscale 16 (2024) 13663, https://doi.org/10.1039/d4nr00952e.
[276] T. Hajilounezhad, R. Bao, K. Palaniappan, et al., Predicting carbon nanotube forest attributes and mechanical properties using simulated images and deep learning, Comput. Mater. 7 (2021) 134, https://doi.org/10.1038/s41524-021-00603-8.
[277] M. Meng, Z.S. Qiu, R.Z. Zhong, et al., Adsorption characteristics of supercritical CO_2/CH_4 on different types of coal and a machine learning approach, Chem. Eng. J. 368 (2019) 874, https://doi.org/10.1016/j.cej.2019.03.008.
[278] S. Davodi, H.V. Thanh, D.A. Wood, et al., Machine-learning models to predict hydrogen uptake of porous carbon materials from influential variables, Sep. Purif. Technol. 316 (2023) 123807, https://doi.org/10.1016/j.seppur.2023.123807.
[279] W. Jiang, X.J. Xing, S. Li, et al., Synthesis, characterization and machine learning based performance prediction of straw activated carbon, J. Clean. Prod. 212 (2019) 1210–1223, https://doi.org/10.1016/j.jclepro.2018.12.093.
[280] J. Moreno, A. Bonilla, D.I. Mendoza, et al., Artificial neural network-based surrogate modeling of multi-component dynamic adsorption of heavy metals with a biochar, J. Environ. Chem. Eng. 6 (2018) 5389–5400, https://doi.org/10.1016/j.jece.2018.08.038.
[281] M. Zhou, A. Vassallo, J. Wu, Data-driven approach to understanding the in-operando performance of heteroatom-doped carbon electrodes, ACS Appl. Energy Mater. 3 (6) (2020) 5993–6000, https://doi.org/10.1021/acsaem.0c01059.
[282] H. Iwata, Y. Hayashi, A. Hasegawa, et al., Classification of scanning electron microscope images of pharmaceutical excipients using deep convolutional neural networks with transfer learning, Int. J. Pharm.: X 4 (2022) 100135, https://doi.org/10.1016/j.ijpx.2022.100135.
[283] J.H. Zhou, P. Yang, P.A. Kots, et al., Tuning the reactivity of carbon surfaces with oxygen-containing functional groups, Nat. Commun. 14 (2023) 2293, https://doi.org/10.1038/s41467-023-37962-3.
[284] P. Rajak, A. Krishnamoorthy, A. Mrishra, et al., Predict. Synth. Quantum Mater. probabilistic Reinf. Learn. (2009) 06739, https://doi.org/10.48550/arXiv.2009.06739 (arXiv).
[285] M.K. Alzahrani, A. Shapoval, Z. Chen, et al., Pore-GNN: a graph neural network-based framework for predicting flow properties of porous media from micro-CT images, Adv. Geo-Energy Res. 10 (2023) 39–55, https://doi.org/10.46690/ager.2023.10.05.



Chuang Wang is currently pursuing a Ph.D. degree in Energy and Power Engineering at Shandong University. He earned his bachelor's degree from Harbin University of Science and Technology and his master's degree from Shandong University. His research focuses on machine learning-assisted pore design of carbon materials.

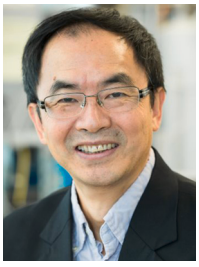
Xingxing Cheng is a professor at the School of Energy and Power Engineering, Shandong University. She obtained her Ph.D. from the University of British Columbia, Canada, and her master's and bachelor's degrees from Southeast University. Her current research interests include carbon materials, hydrogen storage technology, and artificial intelligence/energy storage materials science and technology.

Kai Hong Luo is the Director of the Center for Combustion Energy at Tsinghua University. He earned his Ph.D. from the University of Cambridge, UK. He served as Head of the Energy Systems Division at the University of Southampton in 2004 and at University College London in 2013. He was elected as a Fellow of the Royal Academy of Engineering in 2020 and a Fellow of the European Academy of Sciences in 2022. His current research interests include energy science and technology and high-precision numerical computation.

Krishnaswamy Nandakumar is a professor in the Department of Chemical Engineering at Louisiana State University, USA. He obtained his Ph.D. in Chemical Engineering from Princeton University. He was elected as a Fellow of the Canadian Academy of Engineering in 2007. He previously served as Editor-in-Chief of the International Journal of Nonlinear Sciences & Numerical Simulation and is currently a member of the editorial boards of International Journal of Numerical Analysis & Modeling, Chemical Engineering & Technology, Journal of Chemical and Process Engineering, Journal of Engineering, and Canadian Journal of Chemical Engineering. His research focuses on computer-aided modeling of chemical, mineral, polymer, and electrochemical processes.

Zhiqiang Wang is a professor at the School of Energy and Power Engineering, Shandong University. He earned his Ph.D. from the School of Energy Science and Engineering, Harbin Institute of Technology, and completed postdoctoral research at Shandong University. His current research interests include energy storage and hydrogen energy technology.

Meng Ni is the Associate Dean of the Faculty of Construction and Environment at The Hong Kong Polytechnic University. He obtained his Ph.D. from the Department of Mechanical Engineering, The University of Hong Kong, and his master's and bachelor's degrees from Northwestern Polytechnical University. He is the Co-Editor-in-Chief of Energy Reviews, Senior Editor of Sustainable Energy Technologies and Assessments and e-Prime, and Associate Editor of International Journal of Green Energy, The Canadian Journal of Chemical Engineering, and Energy Storage. His research focuses on hydrogen energy and fuel cells.



Xiaotao Bi is the Director of the Clean Energy Research Center at the University of British Columbia, Canada. He earned his Ph.D. from the University of British Columbia and his master's and bachelor's degrees from Tsinghua University. He completed postdoctoral research at the University of British Columbia and the University of Montreal. He was elected as a Fellow of the Canadian Academy of Engineering in 2014. His current research interests include clean energy technologies and systems in the context of artificial intelligence.

Jiansheng Zhang is a professor in the Department of Thermal Engineering at Tsinghua University. He earned his Ph.D. from Tsinghua University. He served as Deputy Director of the Shanxi Clean Energy Research Institute, Tsinghua University, in 2015. He was elected as a Foreign Member of the Russian Academy of Engineering in 2022. His research focuses on coal conversion and carbon material utilization technologies.

Chunbo Wang is a professor in the Department of Power Engineering at North China Electric Power University. He earned his Ph.D. from the School of Energy and Environment, Southeast University. He worked as a senior visiting scholar at CanmetENERGY, Canada, in 2006 and 2014. His current research interests include carbon capture and high-value utilization.