Machine Learning Prediction of Radioactive Iodine Adsorption by Metal-Organic Frameworks for Nuclear Medicine Wastewater

Authors: Dr. Chunhui Gong, Ming-Jian Sang, Gong, Mrs. Chunhui, Wang, Shu-Ting, Gu, Dr. Ao-Tian, Wang, Dr. Peng, Yang, Dr. Yi, Gong, Mrs. Chunhui

Date: 2025-12-01T00:00:00+00:00

Abstract

The efficient removal of radioactive iodine from nuclear medicine wastewater is of great significance for radiation protection. Metal-organic frameworks (MOFs), owing to their tunable structural characteristics, have demonstrated considerable potential for iodine adsorption. In this study, a dataset was constructed from published literature, and machine learning methods were applied to predict the iodine adsorption capacity of MOFs and to identify the key influencing factors. The results indicated that the XGBoost model exhibited the best predictive performance ($R^2 = 0.96$). Feature importance and SHAP analyses revealed that the type of metal ion exerted the most significant influence, while higher synthesis temperatures and longer reaction times favored the development of crystalline structures and pore channels. In addition, higher drying temperatures effectively activated adsorption sites, thereby markedly enhancing the adsorption capacity. This study provides data-driven support for the rational design of high-performance radioactive iodine adsorbents, offering valuable insights into the treatment of nuclear medicine wastewater and radiation protection.

Full Text

Machine Learning Prediction of Radioactive Iodine Adsorption by Metal-Organic Frameworks for Nuclear Medicine Wastewater

Ming-Jian Sang[1],#, Chun-Hui Gong[1],#,, *Shu-Ting Wang[1], Ao-Tian Gu[1], Peng Wang[1], Yi Yang[1],*

[1] School of Environmental and Biological Engineering, Nanjing University of Science and Technology, Nanjing, 210094, China

# These authors contributed equally to this work.

- Correspondence: gongchunhui@njust.edu.cn; yangyi@njust.edu.cn

Abstract

The efficient removal of radioactive iodine from nuclear medicine wastewater is of great significance for radiation protection. Metal-organic frameworks (MOFs), owing to their tunable structural characteristics, have demonstrated considerable potential for iodine adsorption. In this study, a dataset was constructed from published literature, and machine learning methods were applied to predict the iodine adsorption capacity of MOFs and to identify the key influencing factors. The results indicated that the XGBoost model exhibited the best predictive performance ($R^2 = 0.96$). Feature importance and SHAP analyses revealed that the type of metal ion exerted the most significant influence, while higher synthesis temperatures and longer reaction times favored the development of crystalline structures and pore channels. In addition, higher drying temperatures effectively activated adsorption sites, thereby markedly enhancing the adsorption capacity. This study provides data-driven support for the rational design of high-performance radioactive iodine adsorbents, offering valuable insights into the treatment of nuclear medicine wastewater and radiation protection.

Keywords: Adsorption capacity; Machine learning; Metal-organic frameworks; Radioactive iodine

1. Introduction

With the rapid development of nuclear medicine, the application of radioactive isotopes in disease diagnosis and therapy has become increasingly widespread [**?**]. Among them, iodine-131, with its short half-life and strong tissue specificity, has been extensively used in radionuclide therapy for thyroid cancer [**?**]. According to the 2024 nuclear medicine survey report by the Chinese Medical Association, approximately 197,766 cases of radioactive iodine treatment are conducted annually in China, with a total annual usage of about 740 TBq of radioactive iodine [**?**, **?**]. During the clinical application of radiopharmaceuticals, wastewater containing radioactive iodine is inevitably generated [**?**]. These effluents typically originate from the synthesis of radiopharmaceuticals, laboratory operations, and patient excretion [**?**]. Due to its high solubility and strong mobility in aqueous environments, radioactive iodine can easily diffuse with wastewater and bioaccumulate through the food chain, leading to internal irradiation hazards [**?**]. Therefore, it is crucial to ensure that the radiation exposure to the public does not exceed the legal limit.

One simple yet effective, though time-consuming, method to reduce the transfer of radioactive materials into the sewage system is to concentrate radioactive effluents in dedicated storage tanks or containers, allow them to decay for ten

half-lives, and then conduct radiation level testing before releasing them once they meet national standards [?, ?]. However, with the increasing use of radionuclide therapy, the decay method is costly, requires significant space and resources, and cannot avoid long-term storage risks. Furthermore, this method depends on human management, making it susceptible to equipment failures or natural disasters, which increases the risk of radiation leakage [?, ?]. In contrast, adsorption methods can quickly remove radioactive iodine from wastewater, reduce space requirements, lower management risks, and do not rely on the decay process, making them an efficient and sustainable wastewater treatment solution [?]. Traditional wet scrubbing and solid adsorption technologies for radioactive iodine treatment often suffer from secondary waste generation and potential pollution [?]. Conventional adsorbents such as activated carbon are further limited by low adsorption capacity and poor regeneration efficiency [?].

In recent years, metal-organic frameworks (MOFs), with their high porosity, large surface area, and tunable pore structures, have emerged as promising candidates for radioactive iodine adsorption [?, ?, ?]. MOFs are assembled from metal centers and organic linkers, and their structures can be functionally tailored through secondary building unit (SBU) design, enabling optimization for adsorption, catalysis, and other applications [?, ?]. By varying organic ligands, metal ions, and their ratios, nearly unlimited MOFs can be synthesized. However, this diversity presents the challenge of efficiently screening materials that meet specific performance requirements [?]. Conventional trial-and-error experiments are time-consuming and labor-intensive, while systematic studies on synthesis-structure-performance relationships remain limited [?]. As of September 2022, more than 110,000 MOFs had been deposited in the Cambridge Structural Database, far exceeding the capacity of traditional experimental or computational screening methods [?].

The rapid rise of artificial intelligence has introduced new approaches that connect data science with chemistry. Machine learning, in particular, offers strong capabilities for analyzing high-dimensional data and handling outliers, making it an efficient tool for evaluating large, heterogeneous datasets, optimizing experimental workflows, and establishing structure-property relationships [?]. It has already shown clear advantages in pollutant adsorption prediction, materials design, and parameter optimization. Several studies have demonstrated its utility in high-throughput screening, force field optimization, and performance prediction, successfully guiding experimental synthesis and validation. Collectively, machine learning has greatly accelerated MOF development and provides strong support for the efficient capture and separation of pollutants such as radioactive iodine [?, ?].

2.1 Data Collection and Preprocessing

In this study, the keywords “metal-organic frameworks”, “iodide”, and “adsorption” were used to search the Web of Science, yielding 113 publications. After manual and AI-assisted screening, 23 publications were selected to construct

a dataset containing 56 MOFs and 248 adsorption capacity records, extracted directly from original sources. The dataset included four groups of variables: synthesis parameters, physicochemical properties, adsorption conditions, and adsorption capacity. Synthesis parameters covered factors such as metal and ligand type, mass ratio (M/O), solvent, modification, pyrolysis, reaction conditions, and drying parameters, with categorical and continuous variables appropriately defined. Physicochemical properties included pore volume, pore size, and specific surface area, while adsorption conditions included concentration, dosage, pH, temperature, and time. The output variable was the Langmuir maximum adsorption capacity (Qm) [**?**]. To balance data integrity and representativeness, features with less than 30% missing values were retained and imputed using a combination of simple methods (mean, median, and mode imputation) and advanced techniques (KNN, random forest, and hot-deck), thereby improving data quality and supporting reliable machine learning modeling [**?**, **?**].

2.2 Methods of Feature Engineering

Feature selection is critical for improving the predictive performance and generalization of machine learning models. Strong correlations among features may cause redundancy, bias importance assessment, and reduce interpretability. To address these issues, Spearman correlation coefficients and hierarchical clustering were applied to analyze feature relationships in this study [**?**]. Spearman correlation, a non-parametric measure, effectively captures monotonic relationships in nonlinear or non-normal data. A correlation matrix was constructed to identify strongly related features, and hierarchical clustering with average linkage was used to merge similar ones, preserving key information while reducing feature numbers and mitigating multicollinearity [**?**]. Subsequently, additional screening combined correlation strength with physicochemical relevance to remove redundant or weakly explanatory variables. The final feature set accurately represented the target variable and enhanced model stability and generalizability, providing reliable data support for subsequent modeling [**?**].

2.3 Methods for Constructing Machine Learning Models

The machine learning workflow in this study involved reading the dataset, specifying categorical and continuous variables, defining input and output features, splitting the dataset, configuring models, setting hyperparameters, performing five-fold cross-validation, training the final model, and evaluating accuracy and generalization [**?**]. As hyperparameters strongly influence predictive performance, joint optimization was conducted to ensure fair comparison and prevent overfitting or underfitting. Four ensemble learning models were tested: LightGBM, XGBoost, CatBoost, and Random Forest [**?**, **?**, **?**, **?**]. For Random Forest, optimized hyperparameters included minimum samples split, maximum depth, and number of estimators, while for CatBoost, LightGBM, and XGBoost, they included learning rate, maximum depth, and number of estimators. Model performance was assessed using the coefficient of determination (R^2) and root

mean square error (RMSE) [**?**]. R^2 reflects explanatory power, with values closer to 1 indicating better fit, whereas RMSE measures the average deviation between predicted and actual values, with smaller values indicating higher accuracy. During tuning, maximizing R^2 and minimizing RMSE were set as optimization objectives, and the optimal hyperparameter configurations for each model were determined [**?**, **?**].

2.4 Analysis Methods of Machine Learning Models

Most machine learning algorithms have traditionally been regarded as black-box models, meaning that only the inputs and outputs are observable, while the internal operations and components remain unknown. The internal structure and working principles of such systems are often not fully understood. However, in this study, uncovering the potential relationships between prediction results and different input features was one of the primary objectives. Mainstream machine learning interpretation methods can be broadly categorized into global and local approaches [**?**]. In addition to feature importance analysis, which quantifies the relative weight of each input feature on the iodine adsorption capacity of MOFs, this study employed several advanced interpretability techniques. The Shapley additive explanation (SHAP) method, based on game theory, was used to assign the contribution of each feature to individual predictions. Individual conditional expectation (ICE) analysis was applied to visualize the effect of varying a single feature while holding others constant. Furthermore, partial dependence plots (PDPs) were utilized to quantitatively analyze the influence trends of different features on adsorption capacity and to further explore the key mechanisms of MOFs in the adsorption process [**?**].

3.1 Dataset Analysis and Optimization Process

During dataset collection, 56 MOFs and 248 adsorption capacity records were obtained, comprising 24 input features and one output feature. Four features —pyrolysis atmosphere, heating rate, pyrolysis time, and pyrolysis temperature —had over 60% missing values and were removed to avoid noise. Using a 30% threshold, the average pore size feature was also eliminated. As shown in Table 1 , some input features were categorical variables, such as metal type, ligand type, solvent type, modified and heated features. Fig.1 compares six imputation methods for handling missing values: mean, median, mode, KNN, hot-deck, and random forest. In the violin plots, the height and width represent the data range and distribution density, respectively; the white dot indicates the median, the thick gray line represents the interquartile range, and the thin gray line shows the 95% confidence interval. Missing values were imputed and the effects were evaluated using violin plots. Results showed that, except for total pore volume, the distributions after imputation with all six methods were consistent with the original data. Consequently, hot-deck imputation was applied for total pore volume, while mode imputation was used for other continuous variables. The original data distributions indicated that some process parameters, such

as reaction temperature, reaction time, and pH, were relatively concentrated, whereas mass ratio, drying time, and synthesis temperature exhibited greater dispersion [**?**]. For categorical variables, metal ion type, organic ligand, and solvent type were encoded using label encoding [**?**]. Due to the limited sample size, pyrolysis and modification conditions were binarized, with "1" indicating presence and "0"absence. After applying mode imputation, hot-deck imputation, and label encoding, the dataset was fully cleaned and prepared for subsequent modeling.

3.2 Analysis of Feature Engineering Processing Results

In this study, Spearman correlation matrices and hierarchical clustering were applied to analyze and filter the input features of MOFs. As shown in Fig.2 and Fig.3, the mass ratio of metal ions to organic ligands was significantly positively correlated with specific surface area (0.37, $p < 0.001$) and total pore volume (0.45, $p < 0.001$), indicating that a higher metal ratio favors the formation of well-ordered pore structures [**?**]. Synthesis reaction time was negatively correlated with specific surface area (-0.29, $p < 0.001$), which may be attributed to excessive crystal growth reducing surface area. Drying temperature was positively correlated with both total pore volume and specific surface area ($p < 0.001$), as higher temperatures facilitate solvent evaporation and pore opening [**?**]. Pyrolysis was negatively correlated with total pore volume (-0.35, $p < 0.001$), suggesting that pyrolysis may cause structural collapse. Modification was excluded because it showed no significant correlation with physicochemical properties and the dataset was limited [**?**]. The type of organic ligand was strongly correlated with solvent type (rank correlation coefficient 0.8). To avoid multicollinearity, the more decisive variable, organic ligand type, was retained, while solvent type was removed [**?**, **?**]. Hierarchical clustering further indicated that drying time conveyed similar information to solvent type. Combined with literature findings that drying time had minimal influence, it was also excluded [**?**]. Ultimately, 14 features were retained: metal type, ligand type, mass ratio, hydrothermal reaction time and temperature, drying temperature, heated, total pore volume, specific surface area, concentration, dosage, pH, reaction time and temperature.

3.3 Model Construction and Hyperparameter Optimization

The machine learning model constructed in this study was carried out on the Jutai parallel computing server, with the specific hardware and software configuration summarized in Table 2 . The optimal hyperparameter settings varied across models due to differences in data distribution, feature dimensionality, and algorithmic mechanisms. To ensure reproducibility and rigor, this study referred to parameter-setting approaches reported in the literature and combined them with the characteristics of the present dataset to establish appropriate search ranges [**?**]. Table 3 summarizes the optimal hyperparameter values for each model under their best-performing conditions.

In the Random Forest model, grid search was conducted with maximum depths of 1–7, 100–1200 trees, and minimum split sizes of 2–10. The results (Fig.4.A-B) showed strong parameter dependence: greater tree depth steadily improved R^2 and reduced RMSE, while increasing the number of trees enhanced model stability and accuracy. Minimum split size had limited influence, though smaller values (2 or 4) provided slight gains, suggesting relaxed splitting constraints can capture finer details. For CatBoost, the grid search tested learning rates of 0.005–0.025, depths of 1–10, and 100–1500 iterations. As tree depth and iterations increased, R^2 improved and RMSE decreased (Fig.4.C-D). Performance gains were strongest at low learning rates, where additional iterations compensated for weak individual learners, highlighting CatBoost' s reliance on careful learning rate adjustment and iterative training. LightGBM (Fig.4.E-F) and XGBoost (Fig.4.G-H) shared the same hyperparameter ranges. LightGBM performed best with low learning rates, deeper trees, and more iterations, showing steady accuracy gains. XGBoost showed the sharpest improvement: R^2 was initially negative, reflecting underfitting, but rose rapidly with more iterations, while RMSE dropped. These results confirm that deeper and longer training significantly improves generalization and fitting performance.

3.4 The Training Effect of Machine Learning Models

Fig.5 presents the performance of four machine learning models in predicting the adsorption capacity of MOFs. Red points represent predictions on the training set, and blue points represent predictions on the test set. The gray dashed line denotes the ideal prediction line (predicted value = actual value), and the marginal curves illustrate the data distribution densities of both sets. Overall, all four models demonstrated strong fitting ability, though there were notable differences in performance. The XGBoost model (Fig.5D) achieved outstanding accuracy, with an R^2 of 0.9966 and an RMSE of 12.2583 on the training set, and an R^2 of 0.9620 and an RMSE of 20.8288 on the test set, reflecting its excellent generalization capability. In contrast, the CatBoost model (Fig.5B) performed well on the training set (R^2 = 0.9802, RMSE = 29.7509), but its performance decreased on the test set (R^2 = 0.8701, RMSE = 38.5160), suggesting overfitting. Both Random Forest (Fig 5A [Figure 5: see original paper]) and LightGBM (Fig.5C) achieved R^2 values around 0.93 on the training set. However, LightGBM outperformed Random Forest on the test set (R^2 = 0.9307 vs. 0.8811) and showed slightly lower RMSE (41.8123 vs. 40.5665). This indicates that LightGBM is more adaptable to new data. In summary, XGBoost not only achieved the lowest error and highest fit on the training set, but also maintained the most stable performance on the test set, surpassing the other three models. Therefore, XGBoost was identified as the most suitable model for predicting the adsorption capacity of MOFs, offering a balance of accuracy and generalization.

3.5.1 Input Feature Importance and SHAP Analysis

Fig.6A presents the results of the feature importance analysis, showing that the different categories of input variables contributed unequally to the model. Reaction condition variables accounted for 91.0% of the total importance, far exceeding physicochemical properties (6.1%) and adsorption conditions (2.9%). These results highlight the decisive role of synthesis parameters in determining the structural and functional attributes of MOFs [**?**]. The type of metal ion was the most influential feature at 59.8%, which is consistent with its fundamental role of metal nodes in defining framework topology, surface charge distribution, active site types, and thermal stability [**?**]. For instance, Zn- and Zr-based MOFs exhibit strong affinity for anionic pollutants, such as Cr(VI), whereas Cu- and Fe-based MOFs bind more readily to cationic pollutants. The different valence states of metal ions alter the strength of coordination bonds, which affects framework flexibility and adsorption selectivity [**?**]. Reaction time (15.5%) and drying temperature (5.8%) also ranked highly because reaction time influences crystal development and defect density, and drying temperature impacts pore wall structures and functional group stability. These factors indirectly regulate adsorption by activating MOFs [**?**, **?**, **?**]. Among the adsorption variables, the initial concentration contributed 5.2%, which is higher than the contributions of pH (0.5%) and reaction temperature (2.3%). These results suggest that pollutant concentration gradients significantly impact loading capacity and adsorption saturation, which is consistent with previous experimental findings on adsorption driving forces [**?**]. Of the physicochemical properties, only 2.2% was contributed by specific surface area, suggesting that adsorption capacity is not solely determined by surface area, but rather by pore accessibility and the distribution of functional groups. For example, studies on Zr-MOFs for As(V) adsorption have shown that a higher surface area does not necessarily yield greater adsorption, whereas the number of active sites and electronic density distribution play a more decisive role [**?**].

To further understand the positive and negative influence trends of each variable on the prediction results, a SHAP value analysis was conducted [**?**]. As shown in Fig.6B, the concentration exhibited the widest SHAP value distribution across all samples, indicating its strong impact on model outputs. Generally, higher SHAP values were associated with greater predicted adsorption capacities. This corroborates the feature importance ranking and confirms that different metal ions may cause either positive or negative shifts in model outputs. In addition, synthesis reaction time, drying temperature, and solution concentration also demonstrated clear influence directions. Longer reaction times and moderate drying temperatures were favorable for improving adsorption capacity. Conversely, variables such as pH and adsorbent dosage displayed relatively narrow SHAP distributions, suggesting weaker explanatory power in most samples or localized effects within specific subsets [**?**].

3.5.2 Analysis of Individual Conditional Expectations and Partial Dependencies

The Individual Conditional Expectation (ICE) method was applied to continuous variables in order to analyze sample-level responses and further interpret the mechanisms that influence MOF adsorption performance (Fig.7). Each blue line represents the dependency of a single sample, and the red line indicates the average trend. Consistent with the results of the feature importance and SHAP analyses, the ICE plots revealed heterogeneity in the direction of effects across the samples, reflecting the nonlinear and complex behavior of the MOFs. The metal-to-ligand mass ratio (Fig.7A) showed a sharp increase near ~2.0, suggesting that sufficient metal sites are necessary for promoting framework integrity and generating active sites. Reaction temperature (Fig.7B) showed stable overall effects, though some positive responses were observed around 150–200 °C [**?**, **?**]. Reaction time (Fig.7C) had stronger effects below 10 hours but weakened later, highlighting its importance in early crystal formation. Drying temperature (Fig.7D) and total pore volume (Fig.7E) enhanced adsorption in certain ranges: 50-75 °C for drying and 0.5-1.0 cm^3/g for pore volume, suggesting structural sensitivity zones. Specific surface area (Fig.7F) and initial concentration (Fig.7G) showed dispersed patterns, indicating that heterogeneous effects are influenced by microstructure and diffusion. Below 1000 m^2/g, adsorption increased sharply, but gains plateaued at higher values, reflecting diminishing returns [**?**]. Adsorption temperature (Fig.7J) improved performance above 25 °C, consistent with endothermic adsorption. Low pH (Fig.7H) promoted adsorption by enhancing surface protonation, while adsorbent dosage (Fig.7I) and adsorption time (Fig.7K) showed minimal influence [**?**, **?**].

The partial dependence analysis of continuous features with adsorption capacity is shown in Fig.8. Fig.8A highlighted the interaction between drying temperature and specific surface area. Adsorption capacity increased significantly when the drying temperature was 40–80 °C and the BET was below 1500 m^2/g, after which gains saturated. The effect of drying temperature showed a rise-then-fall trend, suggesting that appropriate drying improved pore utilization. Fig.8B indicated a strong synergy between the drying temperature and the metal-to-ligand mass ratio, where higher ratios (>2) allowed temperature increases to enhance adsorption by providing more coordination centers and improving pore stability. Fig.8C showed adsorption rose markedly when reaction time exceeded 50 h with drying around 140 °C, demonstrating that crystal development required matched drying conditions. Fig.8D-G showed that the specific surface area was most responsive within 70–150 °C synthesis temperatures, while higher synthesis temperatures (>180 °C) combined with drying further improved stability and adsorption. Figures 8H-I revealed boundary effects. When BET exceeded 2000 m^2/g, longer synthesis times or higher temperatures produced only saturated gains, suggesting structural limits. Overall, interactions among the metal-to-ligand ratio, drying temperature, synthesis temperature, and specific surface area were dominant in regulating MOF adsorption capacity, establishing

a mechanistic basis for multiparameter optimization.

Conclusions

This study used machine learning to optimize the synthesis of MOFs for removing radioactive iodine from nuclear medicine wastewater. A dataset of 56 MOFs was constructed, and data cleaning with feature engineering effectively addressed multicollinearity. Four models were developed and tuned via grid search, with results showing that XGBoost achieved the best performance (R^2 = 0.962, RMSE = 20.83). Model interpretation further revealed relationships between key input features and adsorption capacity, providing strong decision support for MOF design and optimization.

However, this study has limitations. The dataset size was relatively small, which could not fully capture the structural diversity of MOFs. Additionally, some microscopic factors influencing adsorption, such as functional site distribution and crystal orientation, were not included as input variables. Future work should expand the dataset with more experimental results to improve model depth and generalization, and integrate deep learning methods to capture nonlinear structure-performance relationships. Combining these approaches with quantum chemical calculations for high-throughput screening will enable a closed-loop framework of modeling, prediction, and optimization, thereby accelerating the engineering application of MOFs in nuclear medicine wastewater treatment and contributing to radiation protection and public health.

Availability of Data

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

Conflict of Interest Statement

All authors disclosed no relevant relationships.

Acknowledgements

This work was partially funded by the National Natural Science Foundation of China (No.52470120 and No.12475310), the China Postdoctoral Science Foundation (2024M764228), and the Fundamental Research Funds for the Central Universities (30925010402).

References

Nagai Y. Production scheme for diagnostic-therapeutic radioisotopes by accelerator neutrons. Proceedings of the Japan Academy, Series B. 2021;97(6):292-323. http://doi.org/10.2183/pjab.97.017

Mihailović J. Evolving Paradigm in Radioactive Iodine Therapy for Differentiated Thyroid Cancer: Historical Perspectives, Current Practices and Future Directions. Diagnostics (Basel, Switzerland). 2025;15(11). http://doi.org/10.3390/diagnostics15111438

Zhang H, Zheng J, Hu N et al. The status of nuclear medicine in China: the first official national survey. European Journal of Nuclear Medicine and Molecular Imaging. 2024;51(8):2172-2178. http://doi.org/10.1007/s00259-024-06687-w

Wang J, Lan X, Shi H et al. A Brief Report on the Results of the 2024 National Survey of Nuclear Medicine Conducted by the Chinese Society of Nuclear Medicine. Journal of Nuclear Medicine 2025;66(5):684-685. http://doi.org/10.2967/jnumed.124.269044

Mu W, Yu Q, Li X et al. Adsorption of radioactive iodine on surfactant-modified sodium niobate. RSC Advances. 2016;6(85):81719-81725. http://doi.org/10.1039/C6RA18091

Ren H, Li Y, Xu L et al. A review of recent advances in health risks and detection of 129I and 131I in environmental and biological samples. Environmental Technology & Innovation. 2025;38:104146. https://doi.org/10.1016/j.eti.2025.104146

Whetter JN, Śmiłowicz D, Boros E. Exploring Aqueous Coordination Chemistry of Highly Lewis Acidic Metals with Emerging Isotopes for Nuclear Medicine. Accounts of Chemical Research. 2024;57(6):933-944. http://doi.org/10.1021/acs.accounts.3c00781

Ma H, Shen M, Tong Y et al. Radioactive Wastewater Treatment Technologies: A Review. 2023;28(4):1935. https://doi.org/10.3390/molecules28041935

Zhang X, Gu P, Liu Y. Decontamination of radioactive wastewater: State of the art and challenges forward. Chemosphere. 2019;215:543-553. http://doi.org/10.1016/j.chemosphere.2018.10.029

Kvashnina K, Claret F, Clavier N et al. Long-term, sustainable solutions to radioactive waste management. Scientific Reports. 2024;14(1):5907. http://doi.org/10.1038/s41598-024-55911-y

Smith J, Marks N, Irwin T. The risks of radioactive wastewater release. Science (New York, NY). 2023;382(6666):31-33. http://doi.org/doi:10.1126/science.adi5446

Zhou X, Wang X, Sun S et al. Synthesis of copper-doped ZIF-8 for enhanced uranium(VI) adsorption: performance and mechanism. Separation and Purification Technology. 2025;378:134705. http://doi.org/10.1016/j.seppur.2025.134705

Lin J, Liu Z, Guo Y et al. Machine learning accelerates the investigation of targeted MOFs: Performance prediction, rational design and intelligent synthesis. Nano Today. 2023;49:101802. http://doi.org/10.1016/j.nantod.2023.101802

Takahashi M, Linh LK, A MS et al. Non-Invasive Prenatal Testing (NIPT) Implementation in Japan: A Comparison with the United Kingdom, Germany,

Italy, Sweden, and Taiwan. International journal of environmental research and public health. 2022;19(24). http://doi.org/10.3390/ijerph192416404

Zhang X, Maddock J, Nenoff TM et al. Adsorption of iodine in metal–organic framework materials. Chemical Society Reviews. 2022;51(8):3243-3262. http://doi.org/10.1039/D0CS01192D

Li J, Wang X, Zhao G et al. Metal–organic framework-based materials: superior adsorbents for the capture of toxic and radioactive metal ions. Chemical Society Reviews. 2018;47(7):2322-2356. http://doi.org/10.1039/C7CS00543A

Yaghi OM, O' Keeffe M, Ockwig NW et al. Reticular synthesis and the design of new materials. Nature. 2003;423(6941):705-714. http://doi.org/10.1038/nature01650

Kim J, Chen B, Reineke TM et al. Assembly of Metal−Organic Frameworks from Large Organic and Inorganic Secondary Building Units: New Examples and Simplifying Principles for Complex Structures. Journal of the American Chemical Society. 2001;123(34):8239-8247. http://doi.org/10.1021/ja010825o

Wang Z, Wang C, Hu Q et al. Covalent-Metal organic Frameworks: Preparation and applications. Chemical Engineering Journal. 2024;483:149217. http://doi.org/10.1016/j.cej.

Tan H, Teng Y, Shan G. High throughput computational screening and interpretable machine learning for iodine capture of metal-organic frameworks. npj Computational Materials. 2025;11(1):115. http://doi.org/10.1038/s41524-025-01617-2

Abdi J, Hadipoor M, Hadavimoghaddam F et al. Estimation of tetracycline antibiotic photodegradation from wastewater by heterogeneous metal-organic frameworks photocatalysts. Chemosphere. 2022;287:132135. http://doi.org/10.1016/j.chemosphere.2021.132135

Auslander N, Gussow AB, Koonin EV. Incorporating Machine Learning into Established Bioinformatics Frameworks. 2021;22(6):2903. https://doi.org/10.3390/ijms22062903

Li S, Zhang Y, Hu Y et al. Predicting metal-organic frameworks as catalysts to fix carbon dioxide to cyclic carbonate by machine learning. Journal of Materiomics. 2021;7(5):1029-1038. http://doi.org/10.1016/j.jmat.2021.02.005

Korolev VV, Nevolin YM, Manz TA et al. Parametrization of Nonbonded Force Field Terms for Metal–Organic Frameworks Using Machine Learning Approach. Journal of Chemical Information and Modeling. 2021;61(12):5774-5784. http://doi.org/10.1021/acs.jcim.1c01124

Liu T, Chen Z, Li Z et al. Preparation of magnetic hydrochar derived from iron-rich Phytolacca acinosa Roxb. for Cd removal. Science of The Total Environment. 2021;769:145159. http://doi.org/10.1016/j.scitotenv.2021.145159

Anderson MJ. Permutational Multivariate Analysis of Variance (PERMANOVA). In: Wiley StatsRef: Statistics Reference Online.1-15. https://doi.org/10.1002/9781118445112.stat0

Schmitt P, Mandel J, Guedj MJJob et al. A Comparison of Six Methods for Missing Data Imputation. 2015;6:0-0. https://doi.org/10.4172/2155-6180.1000224

Liu J, Zhong W, Li R. A selective overview of feature screening for ultrahigh-dimensional data. Science China Mathematics. 2015;58(10):1-22. http://doi.org/10.1007/s11425-0

Cai J, Luo J, Wang S et al. Feature selection in machine learning: A new perspective. Neurocomputing. 2018;300:70-79. http://doi.org/10.1016/j.neucom.2017.11.077

Palansooriya KN, Li J, Dissanayake PD et al. Prediction of Soil Heavy Metal Immobilization by Biochar Using Machine Learning. Environmental Science & Technology. 2022;56(7):4187-4198. http://doi.org/10.1021/acs.est.1c08302

Wang Y, Jian W, Yuan Z et al. Deep learning with attention modules and residual transformations improves hepatocellular carcinoma (HCC) differentiation using multiphase CT. 2025;9(1):13-22. http://doi.org/10.1002/pro6.70003

Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32. http://doi.org/10.1023/A:1010933404324

Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. Journal of Big Data. 2020;7(1):94. http://doi.org/10.1186/s40537-020-00369-8

Sagi O, Rokach L. Approximating XGBoost with an interpretable decision tree. Information Sciences. 2021;572:522-542. http://doi.org/10.1016/j.ins.2021.05.055

Ni X, Wang H, Chen L et al. Classification of Aviation Incident Causes using LGBM with Improved Cross-Validation. Journal of Systems Engineering and Electronics. 2024;35(2):396-405. http://doi.org/10.23919/JSEE.2024.000035

Liu X, Long Z, Yang L et al. Prediction of glass forming ability in amorphous alloys based on different machine learning algorithms. Journal of Non-Crystalline Solids. 2021;570:121000. http://doi.org/10.1016/j.jnoncrysol.2021.121000

Ünalan S, Günay O, Akkurt I et al. A comparative study on breast cancer classification with stratified shuffle split and K-fold cross validation via ensembled machine learning. Journal of Radiation Research and Applied Sciences. 2024;17(4):101080. http://doi.org/10.1016/j.jrras.2024.101080

Shao W, Lin X, Huang Y et al. Rapid patient-specific organ dose estimation in computed tomography scans via integration of radiomics features and neural networks. Quantitative imaging in medicine and surgery. 2024;14(10):7379-7391. http://doi.org/10.21037/qims-24-645

Oviedo F, Ferres JL, Buonassisi T et al. Interpretable and Explainable Machine Learning for Materials Science and Chemistry. Accounts of Materials Research. 2022;3(6):597-607. http://doi.org/10.1021/accountsmr.1c00244

Li H, Wang C, Zeng Y et al. Combining Computational Screening and Machine Learning to Predict Metal-Organic Framework Adsorbents and

Membranes for Removing CH(4) or H(2) from Air. Membranes. 2022;12(9). http://doi.org/10.3390/membranes1209083

Qian T, Li J, Deng Y. Pore structure modified diatomite-supported PEG composites for thermal energy storage. Scientific Reports. 2016;6(1):32392. http://doi.org/10.1038/srep3

Jia BB, Zhang ML. Multi-Dimensional Classification via Decomposed Label Encoding. IEEE Transactions on Knowledge and Data Engineering. 2023;35(2):1844-1856. http://doi.org/10.1109/TKDE.2021.3100436

Park MH, Lee YH, Kim HJ et al. Ferroelectricity and antiferroelectricity of doped thin HfO2-based films. Advanced materials (Deerfield Beach, Fla). 2015;27(11):1811-1831. http://doi.org/10.1002/adma.201404531

Demakov PA. Properties of Aliphatic Ligand-Based Metal-Organic Frameworks. Polymers. 2023;15(13). http://doi.org/10.3390/polym15132891

Zeggai Fz, Ait-Touchente Z, Bachari K et al. Investigation of Metal-Organic Frameworks (MOFs): Synthesis, Properties, and Applications - An In-Depth Review. Chemical Physics Impact. 2025;10:100864. http://doi.org/10.1016/j.chphi.2025.100864

Taheri A, Babakhani EG, Towfighi J. Study of synthesis parameters of MIL-53(Al) using experimental design methodology for CO2/CH4 separation. 2018;36(1-2):247-269. http://doi.org/10.1177/0263617416688690

Zhao F, Tang L, Song W et al. Predicting and refining acid modifications of biochar based on machine learning and bibliometric analysis: Specific surface area, average pore size, and total pore volume. Science of The Total Environment. 2024;948:174584. http://doi.org/10.1016/j.scitotenv.2024.174584

Furukawa H, Cordova KE, O' Keeffe M et al. The chemistry and applications of metal-organic frameworks. Science (New York, NY). 2013;341(6149):1230444. http://doi.org/10.1126/science.1230444

Yang X, Wang Y, Zhao C et al. Health risk and disease burden attributable to long-term global fine-mode particles. Chemosphere. 2022;287:132435. http://doi.org/10.1016/j.chemosphere.2021.132435

Liu Y, Liu S, Gonçalves AAS et al. Effect of metal–ligand ratio on the CO2 adsorption properties of Cu–BTC metal–organic frameworks. RSC Advances. 2018;8(62):35551-35556. http://doi.org/10.1039/C8RA07774F

Kebede TG, Dube S, Nindi MM. Biopolymer electrospun nanofibres for the adsorption of pharmaceuticals from water systems. Journal of Environmental Chemical Engineering. 2019;7(5):103330. http://doi.org/10.1016/j.jece.2019.103330

Gong C-H, Li Z-Y, Chen K-W et al. Synthesis and characterization of Ag@Cu-based MOFs as efficient adsorbents for iodine anions removal from aqueous solutions. Journal of Environmental Radioactivity. 2023;265:107211. http://doi.org/10.1016/j.jenvrad.2023.

Daliran S, Oveisi AR, Kung C-W et al. Defect-enabling zirconium-based metal–organic frameworks for energy and environmental remediation applications. Chemical Society Reviews. 2024;53(12):6244-6294. http://doi.org/10.1039/D3CS01057K

Wu K-J, Wu C, Fang M et al. Application of metal–organic framework for the adsorption and detection of food contamination. TrAC Trends in Analytical Chemistry. 2021;143:116384. http://doi.org/10.1016/j.trac.2021.116384

Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems; 2017; Long Beach, California, USA. https://doi.org/10.48550/arXiv.1705.07874

Wang AY-T, Murdock RJ, Kauwe SK et al. Machine Learning for Materials Scientists: An Introductory Guide toward Best Practices. Chemistry of Materials. 2020;32(12):4954-4965. http://doi.org/10.1021/acs.chemmater.0c01907

Wang H, Dong X, Lin J et al. Topologically guided tuning of Zr-MOF pore structures for highly selective separation of C6 alkane isomers. Nature Communications. 2018;9(1):1745. http://doi.org/10.1038/s41467-018-04152-5

Sun Y-X, Sun W-Y. Influence of temperature on metal-organic frameworks. Chinese Chemical Letters. 2014;25(6):823-828. http://doi.org/10.1016/j.cclet.2014.04.032

Gong C-H, Li Z-Y, Chen K-W et al. Synthesis and characterization of Cu2O@Cu/Al-C LDH for efficient adsorption of iodide anions in aqueous solutions. Journal of Radioanalytical and Nuclear Chemistry. 2023;332(7):2793-2805. http://doi.org/10.1007/s10967-023-08943-z

Beydaghdari M, Hooriabad Saboor F, Babapoor A et al. Recent Advances in MOF-Based Adsorbents for Dye Removal from the Aquatic Environment. 2022;15(6):2023. https://doi.org/10.3390/en15062023

Ming Y, Purewal J, Liu Da et al. Thermophysical properties of MOF-5 powders. Microporous and Mesoporous Materials. 2014;185:235-244. http://doi.org/10.1016/j.micromeso.2013.11.015

Note: Figure translations are in progress. See original paper for figures.