MetaResource
DOI 10.23919/METAR.2026.000034
Review Article

Machine learning throughout the non-ferrous metal lifecycle: An overview and application advances

Haiyang Liao[1], Lei Huang[1](✉), Zhenxing Wang[2], Zhen Zeng[1], Xue Jia[3], Qisheng Huang[1], Jia Yan[1], Meng Li[1], Hongguo Zhang[1](✉), Zhenxin Chen[1], Hao Li[3](✉)

***MetaResource*, Just Accepted Manuscript** • 10.23919/METAR.2026.000034
https://publish.cnki.net/chain on April 10, 2026

Just Accepted

Machine Learning Throughout the Non-Ferrous Metal Lifecycle: An Overview and Application Advances

Haiyang Liao [a], Lei Huang [a, *], Zhenxing Wang [b], Zhen Zeng [a], Xue Jia [c], Qisheng Huang [a], Jia Yan [a], Meng Li [a], Hongguo Zhang [a, *], Zhenxin Chen [a], Hao Li [c, *]

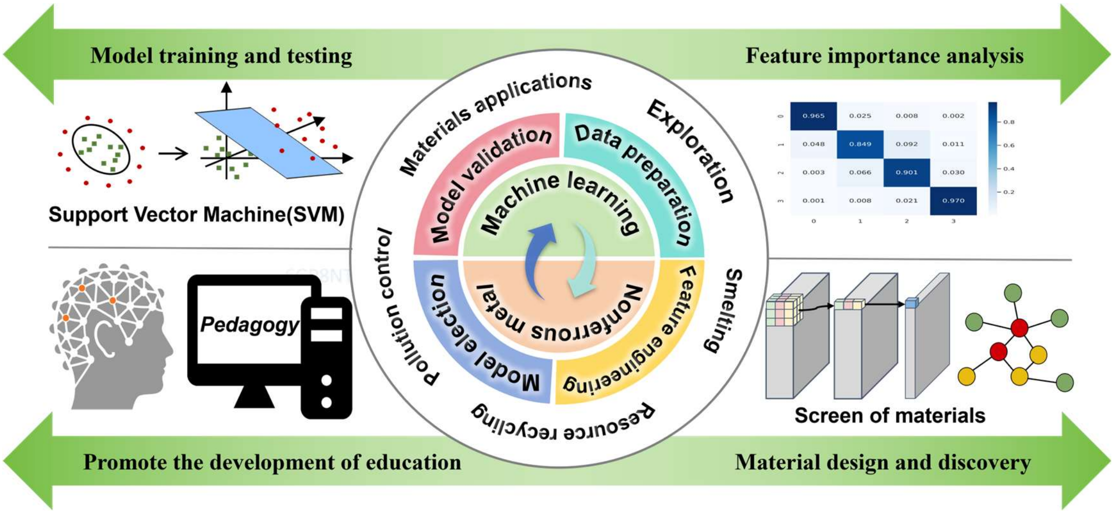
a School of Environmental Science and Engineering, Guangzhou University, Guangzhou 510006, China

b South China Institute of Environmental Sciences, Ministry of Ecology and Environment, Guangzhou 510655, China

c Advanced Institute for Materials Research (WPI-AIMR), Tohoku University, Sendai 980-8577, Japan

* Corresponding author at School of Environmental Science and Engineering, Guangzhou University, Guangzhou Higher Education Mega Center, Guangzhou 510006, China. Advanced Institute for Materials Research (WPI-AIMR), Tohoku University, Sendai 980-8577, Japan. E-mail address: huanglei@gzhu.edu.cn, hgzhang@gzhu.edu.cn, li.hao.b8@tohoku.ac.jp.

Graphical Abstract



Abstract:

Data-driven machine learning (ML) technologies offer significant advantages throughout the entire lifecycle of non-ferrous metal mineral exploration, smelting processes, resource recovery and circular economy, pollution control and environmental remediation, materials applications and development due to their efficiency and robust analytical capabilities. The integration of ML algorithms such as Random Forest (RF), Convolutional Neural Networks (CNN), and XGBoost with reinforcement learning-enabled process control technologies has made substantial contributions to minimizing resource waste. This review summarizes the technological advancements of machine learning techniques in mineral exploration prediction, metallurgical parameter optimization, metal recovery from solid waste, and non-ferrous metal pollution control. Furthermore, we highlight the challenges currently faced in machine learning applications, including data quality and scale issues, the complexity of feature engineering, and poor model interpretability. It also proposes future research priorities focused on optimizing data extraction methods, enhancing model generalization capabilities, and improving interpretability.

Keywords: machine learning; non-ferrous metals; resource recycling; artificial intelligence

1. Introduction

Non-ferrous metals refer to the collective term for metallic elements and their alloys other than ferrous metals, such as iron and manganese, primarily including tungsten, molybdenum, cobalt, vanadium, lead, copper, and others. They are typically categorized into light metals, heavy metals, precious metals, and rare metals. Non-ferrous metals generally exhibit excellent electrical and thermal conductivity[1], corrosion resistance, and wear resistance[2], coupled with outstanding plasticity, ductility, and formability[3]. Consequently, they serve both as structural materials supporting engineering applications and as functional materials meeting specific requirements for electrical conductivity, thermal conductivity, and catalysis. For instance, tungsten has been ideally used as a material for high-power integrated circuits[4], molybdenum-based compounds function in electrochemical capacitors[5], and cobalt-based catalysts activate peroxymonosulfate[6]. At the industrial level, non-ferrous metals are extensively applied in power transmission and distribution equipment, electrical appliances, electronic information devices, and critical defense and aerospace systems.[7,8,9] They form a vital material foundation for ensuring the stable operation of modern industrial systems and driving the development of strategic emerging industries.

However, the large-scale extraction of non-ferrous metal resources simultaneously imposes a significant energy burden and environmental pollution. Data from the International Energy Agency (IEA) indicate that the production and use of aluminum directly generate nearly 270 million tons of carbon dioxide. The mining and smelting processes for non-ferrous metals such as copper, tungsten, lead, and zinc also impose significant energy demands across the industrial sector. Concurrently, the massive emissions of greenhouse gases and toxic substances have severely polluted the atmosphere, water bodies, and soil. Given the non-ferrous metals industry's high energy consumption and pollution levels, compounded by the world's escalating environmental pressures, it is imperative to optimize the entire lifecycle of non-ferrous metals mineral exploration, smelting processes, resource recovery and circular economy, pollution control and environmental remediation, materials applications and development. Traditional methods, often reliant on experience, are inefficient, struggle to address complex issues, and lack systematic approaches.

Due to the advancements in computing power and technological development, machine learning (ML) has emerged as one of the most popular research areas in recent years[10], attracting numerous researchers and practitioners across fields such as finance, medicine, agriculture, and aerospace. As an interdisciplinary field encompassing computer science, mathematics, statistics, and informatics, ML technology continuously learns underlying patterns from multidimensional data, revealing how feature values influence output variables under different conditions. This empowers researchers to conduct efficient, systematic studies while reducing costs and shortening research cycles [11]. Over the past few decades, ML has also gained increasing traction among professionals in the non-ferrous metals industry. Applying ML to this sector optimizes the entire lifecycle of non-ferrous metals—from exploration and smelting to recycling, pollution control, and material applications. Empowering pollution control and resource recovery with data-driven technologies significantly reduces pollution caused by process improvements, lowering both costs and energy consumption.

2. Literature Review and Current Research Status

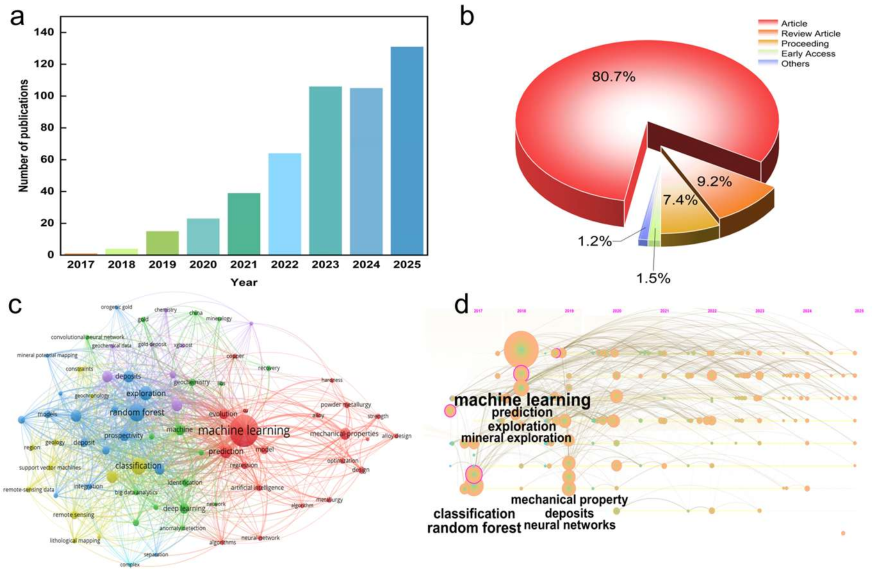


Fig. 1 (a) Annual number of publications from 2017 to 2025. (b) Proportion of different document types (articles, reviews, proceedings, etc.). (c) Keyword co-occurrence network showing major research topics and their relationships. (d) Temporal evolution of high-frequency keywords from 2017 to 2025. Records were retrieved from Web of Science Core Collection using a Topic search (TS): TS=(("machine learning" OR "artificial intelligence" OR "deep learning") AND ("non-ferrous" OR "base metal" OR aluminum OR zinc OR lead OR nickel OR cobalt)), with publication years limited to 2017–2025.

With the rapid advancement of machine learning technology, an increasing number of studies have focused on its applications in the non-ferrous metals industry, particularly in mineral exploration, metallurgical process control, resource recovery, and pollution control. As an effective research method, bibliometric analysis can reveal the research trends and development directions in this field. By analyzing the volume of existing literature, keyword co-occurrence patterns, and author collaboration networks, the research hotspots of machine learning in the non-ferrous metals industry over recent years can be clearly presented. As shown in **Fig. 1(a)** and **Fig. 1(b)**, from 2017 to 2025, the number of publications on machine learning applications in mineral resource prediction has significantly increased, with research papers constituting the vast majority. This indicates a continuous rise in research demand and application potential within this field. Furthermore, Keyword Co-occurrence Network **Fig. 1(c)** and its timeline diagram **Fig. 1(d)** reveal the diverse applications of machine learning in this field, such as the extensive use of Support Vector Machines (SVM), Random Forests (RF), and Convolutional Neural Networks (CNN) in mineral exploration and smelting processes[12]. For instance, Duan et al. successfully removed haze from captured images by employing a hybrid version of VGG16, a lightweight U-Net model proposed in Ref. [13]. Simultaneously, the application of machine learning technologies across the non-ferrous metals industry spans multiple stages, yielding notable achievements particularly in mineral resource prediction, metallurgical process optimization, pollution control, and resource recovery. Research indicates breakthrough progress in applying machine learning algorithms such as SVM, RF, and CNN to mineral exploration. For instance, SVM-based feature selection methods effectively predict mineral resource potential, while convolutional neural networks have been successfully employed for ore type classification. Furthermore, machine learning applications in metallurgical processes, particularly in dynamic optimization control, have yielded remarkable outcomes[14]. By regulating parameters such as temperature and flow rate during smelting, these techniques enhance product quality while reducing energy consumption. Simultaneously, significant advancements have been made in pollutant prediction and waste recovery rate improvement, establishing machine learning as an effective tool for addressing environmental challenges.

Despite significant progress in applying machine learning technologies to the non-ferrous metals industry, numerous challenges remain. First, data quality and volume issues continue to constrain the successful deployment of machine learning models, particularly in mineral exploration and metallurgical process control, where data is often incomplete and noisy, compromising model accuracy and stability[15]. Second, the complexity of feature engineering poses a critical challenge[16]. Extracting effective features from vast amounts of raw data and performing reasonable feature selection and dimensionality reduction remain pivotal steps for successful machine learning implementation. Furthermore, the interpretability of machine learning models presents a significant challenge. Particularly in deep learning models, their “black box” nature leads to poor explainability and traceability of results, posing significant risks in sectors requiring high transparency and regulatory oversight.

This study systematically summarizes the application of machine learning technologies in the non-ferrous metals industry, particularly in mineral exploration, metallurgical processes, resource recovery, and pollution control. By comparing the advantages and disadvantages of different algorithms, this paper reveals how machine learning effectively enhances resource utilization efficiency, reduces energy consumption, and minimizes environmental pollution within the non-ferrous metals sector. Second, we discuss the practical applications of various machine learning algorithms in the non-ferrous metals sector, demonstrating how these techniques address complex challenges beyond the scope of traditional methods. Examples include deep learning models for mineral image classification and ore type identification, alongside instances of improving model accuracy through semi-supervised learning in data-scarce environments. Furthermore, this paper delves into the pivotal role of machine learning algorithms in feature engineering, particularly in data preprocessing, feature selection, and dimensionality reduction techniques. It summarizes successful approaches from existing research and highlights their potential applications within the non-ferrous metals sector. Finally, addressing the challenges of generalization and interpretability in existing machine learning models, this paper analyzes how cross-validation, ensemble learning, and model optimization strategies can enhance model stability and reliability, drawing on the latest research findings. Through this review of methodologies, the paper provides essential theoretical foundations and practical guidance for machine learning applications in the non-ferrous metals industry.

3. Machine Learning Empowering Mining: From Algorithms to Evaluation

3.1. Overall ML workflow in non-ferrous metal research

In recent years, reports integrating ML with the mining industry have proliferated. These studies leverage geological exploration data from diverse sources and metallurgical process parameters, focusing on key characteristic parameters such as mineral chemical composition, ore type, temperature, and pressure. Through systematic training and validation of existing databases using machine learning algorithms such as Support Vector Machines (SVM), Random Forests (RF), and Convolutional Neural Networks (CNN), these studies reveal the critical role of various characteristic parameters in mineral exploration, smelting processes, resource recovery, pollution control, and material applications. They have established a foundational framework for applying ML technologies to the non-ferrous metals industry and achieved effective increases in non-ferrous mineral production.

Applying ML to non-ferrous metal research processes primarily involves data extraction and database establishment, feature engineering, model training and testing, and model application. By collecting, storing, and managing existing feature values and data to form a database, preparations are made for building and training artificial intelligence models. The quality and quantity of the database significantly impact model performance[17]. However, traditional experimental methods are time-consuming and labor-intensive, yielding sparse data, making the collection of accurate and sufficient data a critical step. Feature engineering primarily encompasses data and feature preprocessing, feature selection, dimensionality reduction, and feature combination. Methods such as cleaning, scaling data, filtering, wrapping, and embedding feature selection enable the identification of the most representative feature descriptors from raw data. Techniques such as reducing feature count and feature interaction help avoid the curse of dimensionality while enhancing model generalization performance. Popular ML models include Extreme Gradient Boosting (XGBoost), RF, ANN, SVM, and CNN. Each model has distinct strengths and weaknesses. Cross-validation is essential for evaluating models and algorithms during data processing to select the best-performing prediction model, preventing overfitting or underfitting[18]. Common methods include K-fold cross-validation (K-fold CV)[19], leave-one-out cross-validation (LOOCV)[20], and leave-one-out validation[21]. Cross-validation enables continuous optimization of model performance to meet requirements, ultimately deploying the model in practical applications. Its robust predictive capabilities support research across the entire lifecycle of non-ferrous metals.

Given the scarcity of data in the non-ferrous metals industry, multi-source data and complex feature values are often difficult to systematically collect, posing a challenge for artificial intelligence applications. Some researchers have already initiated such studies. For instance, Shirmard et al. [22] proposed that remote sensing datasets could serve as a novel source of geological exploration data. As shown in **Fig. 2(a)**, machine learning models are trained using remote sensing datasets, utilize the models to predict mineral value in a specific region, and finally validate the predictions through field surveys. Integrating this approach can overcome the limitations of relying solely on field-extracted feature parameters, demonstrating that ML techniques can effectively process complex remote sensing data. According to research by McCoy and Auret [23], feature parameters can be categorized into process measurement data, chemical composition data, mineralogical data, image/video data, equipment status data, and operator input data. They also propose incorporating machine vision technology. Xiong et al. [24] utilized rare event logistic regression (RELR) based on geographic information system (GIS) data to map mineral prospects in southwest Fujian Province, China, with the Receiver Operating Characteristic (ROC) curve shown in **Fig. 2(b)**. However, current ML data primarily originates from experimental research, limiting its industrial application. Nevertheless, these findings provide crucial theoretical guidance for data and feature extraction, successfully reducing both time and economic costs in the non-ferrous metal industry's development.

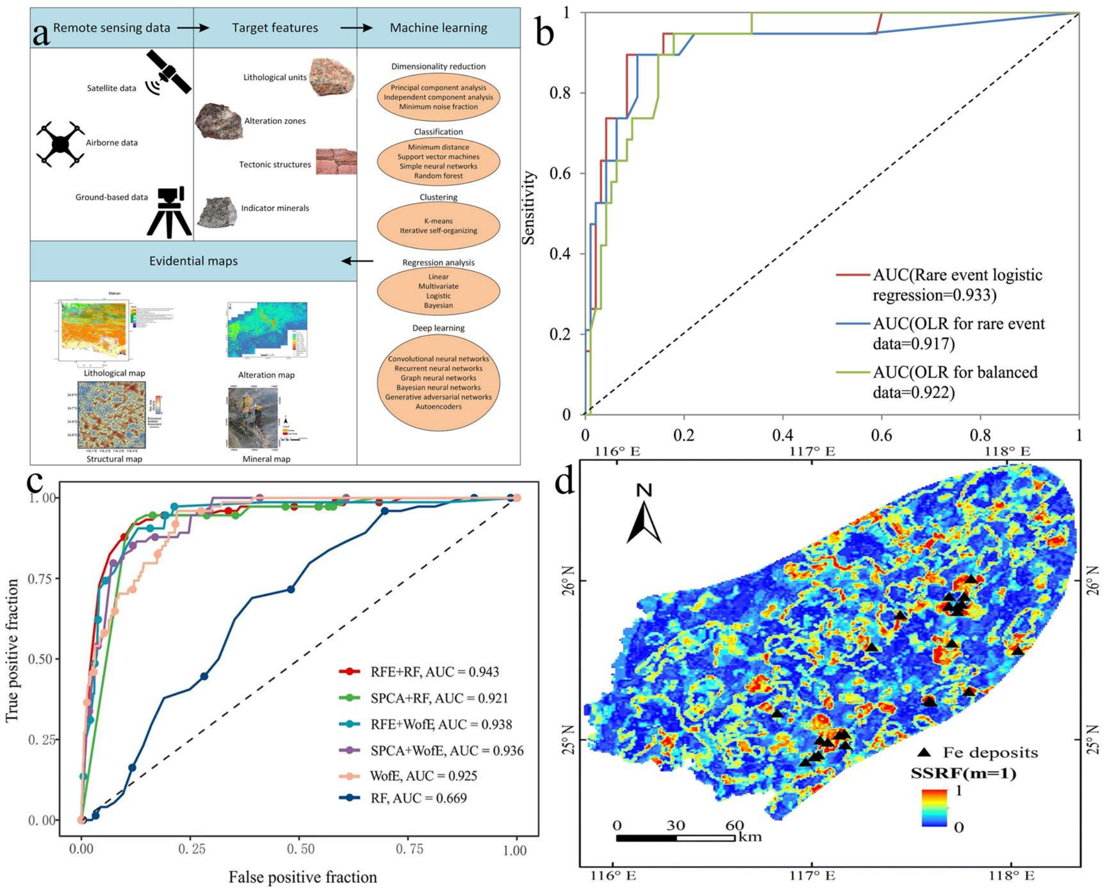


Fig. 2 (a) Workflow for creating evidence maps using remote sensing data and ML methods. Reproduced with permission from Ref. [22]. Copyright 2022, Elsevier. (b) Receiver operating characteristic (ROC) curves for rare event logistic regression and original logistic regression. Reproduced with permission from Ref. [24]. Copyright 2018, Springer Nature. (c) ROC curves of posterior probability maps for mineral prospect predictions generated using four combinations of predictive variables and data integration methods. Reproduced with permission from Ref. [25]. Copyright 2022, Springer Nature. (d) Mineral prospect map derived from semi-supervised random forest (SSRF). Reproduced with permission from Ref. [26]. Copyright 2019, Springer Nature.

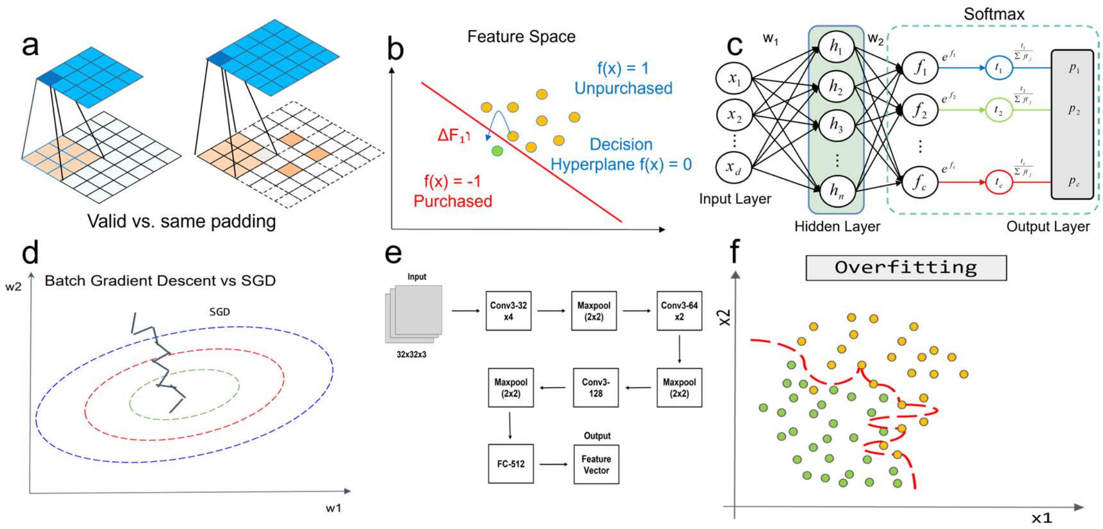


Fig. 3 (a) Illustration of valid vs. "same" padding in 2D convolution. (b) Linear decision boundary f(x)=0 separating two classes (purchased vs. unpurchased samples). (c) Multi-layer perceptron (MLP) with hidden units $\{h_i\}$ and a softmax output layer producing class probabilities $\{p_c\}$. (d) Batch gradient descent versus stochastic gradient descent (SGD) on a convex loss surface. (e) Example convolutional neural network (ConvNet) configuration for feature extraction. (f) Illustration of overfitting.

3.2. Common ML algorithms

Algorithms and models form the core of ML methods, making the construction of an appropriate ML framework crucial. Common ML models can be categorized into supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Among these, supervised learning, unsupervised learning, and semi-supervised learning primarily rely on existing datasets for model training, while reinforcement learning learns and optimizes strategies through interactions between an agent and its environment. Supervised learning requires substantial labeled data and holds significant application value in fields such as non-ferrous metal resource prediction and post-production pollution assessment. A notable example is the innovative feature selection method based on Recursive Feature Elimination (RFE) proposed by Wang et al. [25], which achieves high-precision prediction of mineral resource potential. Notably, their developed RFE-RF hybrid model demonstrated outstanding performance, as shown in **Fig. 2(c)**, achieving an area under the ROC curve (AUC) value of 0.943—the highest among all compared models. Meanwhile, the semi-supervised random forest model proposed by Wang et al. [26] was successfully applied to assess the mineral potential in the southwestern Fujian mineral belt of China, significantly improving the prediction accuracy. The mineral prospect distribution map generated by this model **Fig. 2(d)** visually highlights the prospecting targets within the study area, providing crucial guidance for mineral exploration.

In addition to Random Forests (RF) and Support Vector Machines (SVM), Convolutional Neural Networks (CNN) have also gained prominence in non-ferrous metal research due to their ability to handle spatial data and complex patterns[27]. **Fig. 3(a)** shows the process of convolution and the concept of padding in CNNs. The image illustrates how filters are applied to the input data, and how the output feature maps are generated during the convolution process. CNNs are particularly useful in applications like mineral image classification and defect detection in mining processes. Their ability to extract hierarchical features from raw data makes them ideal for tasks where raw sensor or image data is involved.

Support Vector Machines (SVM) are particularly effective in classification tasks. By constructing an optimal hyperplane to separate different classes, SVM have been widely used in mineral classification and ore type prediction. As discussed in studies by Farhadi et al. [28] and Daviran et al. [29], SVM models have significant applications in both identifying complex geological data and mapping mineral prospects. **Fig. 3(b)** illustrates the decision boundary in a binary classification task. The figure shows how the hyperplane is constructed to separate the two classes, where the decision boundary is at $\mathrm{f}(x) = 0$. Their effectiveness in high-dimensional spaces is one of the reasons why SVM are commonly applied in mining exploration and other predictive tasks.

Multi-layer Perceptron (MLP), a type of artificial neural network (ANN), is another popular supervised learning model that excels in tasks requiring pattern recognition[30]. **Fig. 3(c)** shows the architecture of a multi-layer perceptron (MLP), where the input features $x_1, x_2, \ldots, x_d$ are passed through hidden layers to generate the output predictions. MLPs consist of multiple layers of neurons, allowing for non-linear transformation of input data and complex decision boundaries. These networks are frequently applied to tasks like mineral resource prediction and post-processing quality assessment, where the relationships between features are complex.

Unsupervised learning is commonly used for exploratory analysis of data and features, serving as an effective auxiliary method. For example, as mentioned by Bhat et al.[31], unsupervised learning can provide effective analytical methods for high-dimensional alloy data while revealing structural features and underlying relationships that are difficult to uncover using traditional approaches. Semi-supervised learning can fully utilize limited labeled data and large amounts of unlabeled samples for model training, demonstrating unique advantages in mineral resource prediction. Reinforcement learning enables dynamic selection of optimal strategies for metallurgical processes through autonomous decision-making mechanisms, establishing a well-developed theoretical framework in metallurgical process control. By learning optimal control strategies through interaction with the environment, this method offers innovative solutions for optimizing metallurgical process parameters. For instance, Zheng et al. [32] proposed a hybrid model-driven reinforcement learning (HMBRL) algorithm that *enables intelligent regulation of key parameters such as air volume and tailings flow rate, providing crucial theoretical support and methodological references for automated optimization of complex metallurgical processes. To enhance the efficiency and precision of resource exploration, the hybrid ML model proposed by Tahmasebi et al. holds significant reference value. This model innovatively integrates the strengths of fuzzy logic, neural networks, and genetic algorithms to construct a multi-method fusion intelligent prediction framework, significantly improving the accuracy of mineral resource potential assessment [33]. The rapid advancement of ML technologies presents transformative opportunities for the mining industry. Among these, algorithm optimization and model architecture innovation—as core components of machine learning applications—have emerged as the current research frontiers and key focus areas within the mining sector.

3.3. Model evaluation metrics and validation strategies

To comprehensively evaluate the performance and generalization capabilities of machine learning models, effective validation strategies and evaluation metrics must be employed. Evaluation metrics not only quantify a model's predictive effectiveness for specific tasks but also reflect its performance on unseen data, thereby ensuring reliability. Commonly used metrics include accuracy, precision, recall, F1 score, root mean square error (RMSE), and area under the receiver operating characteristic curve (AUC)[34,35]. Accuracy is the most fundamental and widely applied metric, measuring the proportion of samples correctly predicted by the model. However, accuracy can be misleading when handling imbalanced datasets, particularly when positive and negative categories are unevenly distributed. Nevertheless, when categories are relatively balanced-especially in resource prediction tasks such as mineral resource forecasting-accuracy remains a commonly used and effective indicator. In contrast, precision and recall become particularly crucial for imbalanced data. Precision measures the proportion of all predicted positive samples that are correctly classified, while recall assesses the model's ability to identify positive samples. To holistically evaluate the trade-off between precision and recall, the F1 score—as their harmonic mean—plays a vital role in tasks such as pollution control and resource recovery, where false positives and false negatives can have severe consequences. For regression tasks, RMSE is frequently employed as an evaluation metric, measuring model accuracy by calculating the average of squared errors between predicted and actual values. In tasks such as alloy mechanical property prediction or pollutant concentration estimation, RMSE serves as a critical metric, effectively reflecting a model's performance in continuous numerical forecasting. Finally, AUC is a vital metric for evaluating binary classification models. It measures a model's discrimination capability across different classification thresholds by calculating the area under the ROC curve; a higher AUC value indicates superior classification performance. AUC is widely applied in binary classification tasks such as mineral resource prediction and pollution risk assessment, particularly suited for handling imbalanced datasets. **Fig. 3(d)** compares batch gradient descent and stochastic gradient descent (SGD) during model training, highlighting different convergence behaviors that can affect the final decision scores and, consequently, the ROC curve and AUC, further emphasizing AUC's importance as an evaluation tool for handling complex data.

For model validation strategies, cross-validation, leave-one-out cross-validation, training and testing set partitioning, and bootstrapping are common validation methods. Cross-validation is a technique that divides a dataset into K subsets, where the model is trained on K-1 subsets and tested on the remaining one. This process is repeated K times, each time using a different subset as the test set. The model's performance is ultimately evaluated by averaging the results across all K tests. K-fold cross-validation is particularly effective in assessing a model's stability and robustness. **Fig. 3(e)** illustrates a typical convolutional neural network (CNN) architecture. Such models are prone to overfitting due to their high capacity, making it necessary to adopt appropriate validation strategies to ensure robust generalization performance and reliability. Leave-one-out cross-validation (LOOCV) is a special case of K-fold cross-validation. In this method, one data point is reserved as the test set each time, while the remaining data serve as the training set. Since every data point participates in both training and testing, LOOCV is typically used for smaller datasets, ensuring each sample influences model training and evaluation[36]. Splitting the dataset into training and test sets is the simplest validation method. This approach is suitable for sufficiently large datasets, enabling efficient validation by dividing the data into training and test portions. The validation strategy can typically be selected based on the sample size. For small databases, LOOCV and cross-validation should be prioritized to maximize data utilization. Conversely, large datasets are better suited for validation using a fixed training and testing set partitioning approach. The bootstrap method generates multiple samples by sampling with replacement from the training data. Each model is trained on one bootstrap sample and validated on external data. Bootstrapping is particularly effective for limited datasets, significantly enhancing model stability and generalization capability. **Fig. 3(f)** illustrates the overfitting phenomenon, where a model attains seemingly excellent performance on the training set but generalizes poorly to unseen data, resulting in inflated and unreliable performance estimates. This observation underscores the necessity of employing rigorous validation protocols (e.g., K-fold cross-validation) and appropriate regularization strategies to mitigate overfitting and ensure robust generalization.

Here we delve into several major machine learning models, including supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Supervised learning algorithms demonstrate significant application value in mineral resource prediction and metallurgical process control. Unsupervised and semi-supervised learning are widely used for data analysis and feature extraction, particularly in mineral exploration and resource recovery. Reinforcement learning provides dynamic solutions for optimizing metallurgical processes through interaction with the environment. We also examine common metrics and validation methods for evaluating machine learning models, such as accuracy, precision, recall, F1 score, RMSE, and AUC. These evaluation metrics hold significant importance in fields such as resource prediction and pollution control. Concurrently, validation strategies including cross-validation, leave-one-out cross-validation, and bootstrapping play a crucial role in ensuring model stability, optimizing performance, and preventing overfitting. The effectiveness of different validation methods is demonstrated intuitively. We not only understand the strengths and limitations of common machine learning algorithms but also master how to enhance model performance in the non-ferrous metals industry through effective evaluation and validation strategies. These methods provide theoretical support for the practical application and research of machine learning models, offering crucial guidance for industry professionals in model selection and optimization.

4. Key Applications in Non-Ferrous Metal Lifecycle

With the rapid advancement of artificial intelligence technology, ML has been deeply integrated into every stage of non-ferrous metals' lifecycle management. This encompasses critical domains such as resource exploration prediction, metallurgical process control, resource recycling, environmental monitoring and remediation, as well as material development and application. By intelligently mining multi-source heterogeneous data within the mining industry and systematically optimizing characteristic parameters, this technology not only significantly reduces production costs but also substantially enhances resource utilization efficiency. It provides innovative solutions for advancing the non-ferrous metals industry toward a green and low-carbon transformation.

4.1. Mineral Exploration and Mapping

Resource exploration and prediction, as the initial stage of the non-ferrous metal industry chain, directly impacts the economic benefits of subsequent development through its technological sophistication. With the continuous advancement of machine learning technologies, intelligent algorithmic models—represented by Gradient Boosted Decision Trees (GBDT), Deep Neural Networks (DNN), and SVM—have demonstrated exceptional performance advantages in mineral resource prediction. The GBDT algorithm can enhance predictive performance by building decision tree ensembles through boosting frameworks, while DNNs are multi-layer neural networks capable of recognizing complex patterns and performing feature extraction. [37] These methods not only significantly enhance exploration efficiency but also substantially improve prediction accuracy, offering revolutionary technological supplements to traditional exploration approaches. For instance, Maitre et al. [38] combined optical microscope images with machine vision technology to classify mineral particles such as plagioclase, pyroxene, and manganese pyrolusite. The accuracy of three models—trained Classification and Regression Trees (CART), KNN, and Random Forest (RF)—reached approximately 90%, validating the feasibility of machine learning-based mineral particle recognition methods. Sun et al.[39] emphasized that integrating machine learning methods with mineral geochemistry can help elucidate the genesis mechanisms of lead-zinc (Pb-Zn) ore fields. With the advancement of emerging algorithms such as deep learning, machine learning has not only provided new theoretical frameworks for mineral resource prediction but also spawned a series of innovative models, including deep convolutional neural networks and graph neural networks. These technological advances continue to propel the mineral industry's shift from an experience-driven to a data-driven paradigm, injecting robust momentum into industrial upgrading.

4.2. Smelting and Metallurgical Optimization

Metallurgical processes consume substantial resources, and optimizing their procedures is a critical step toward reducing energy consumption and enhancing production efficiency. To improve energy utilization, researchers have integrated machine learning techniques with alloy properties and metallurgical process parameters, successively developing various predictive models [40] that typically demonstrate high accuracy. For instance, Xiong et al. employed an RF model to predict the hardness and ultimate tensile strength of complex concentrated alloys (CCA), achieving correlation coefficients exceeding 0.9 for both target properties [41]. In Laha et al.'s study [42], RF, ANN, Dynamic Evolutionary Neuro-Fuzzy Inference System (DENFIS), and Support Vector Regression (SVR) were utilized to forecast steel yield during smelting processes. DENFIS is a hybrid system that integrates the adaptive capabilities of evolutionary algorithms with the learning capabilities of neuro-fuzzy systems. **Fig. 4(a)** presents the box plots of the input-output parameters in the model. Researchers selected 10 input variables influencing the steelmaking process and analyzed 54 sets of input-output data patterns to investigate these variables' impact on output yield. **Fig. 4(b)** displays the MSE values of various models, where the support vector regression model and DENFIS achieve training and testing MSE values close to 0.00001. These research outcomes demonstrate outstanding performance in predicting specific parameters. In recent years, improvements in database quality and model performance have significantly increased the utilization rate of non-ferrous metal resources. Concurrently, the introduction of physical metallurgy models and density functional theory has further advanced machine learning techniques. For instance, in 2019, Shen et al. [43] employed physical metallurgy (PM) principles to generate intermediate parameters such as equilibrium volume fractions and precipitation driving forces, then combined these with an SVM model to predict the hardness of ultra-high-strength stainless steel. **Fig. 4(c)** displays the mean and standard deviation of R calculated for 500 partitioned datasets, along with the correlation analysis results. The study revealed that nickel, molybdenum concentration, and aging temperature significantly impact the prediction accuracy, with these variables making substantial contributions to the hardness of R-phase-strengthened ultra-high-strength (UHS) stainless steel. Simultaneously, as shown in **Fig. 4(d)**, 100 different random partitions and multiple-hold methods were employed to randomly divide the training and test sets, investigating the impact of partitioning methods on the generalization capability of the current system. The final R^2 value reached 0.92. The successful integration of physical metallurgical characteristics with machine learning techniques substantially enhances model accuracy and design efficiency. Nevertheless, data scarcity and insufficient model generalization continue to constrain the application of machine learning in metallurgy. Future efforts must focus on optimizing data extraction methods and model development techniques to accelerate the advancement of new metallurgical technologies and advance toward greener metallurgical practices.

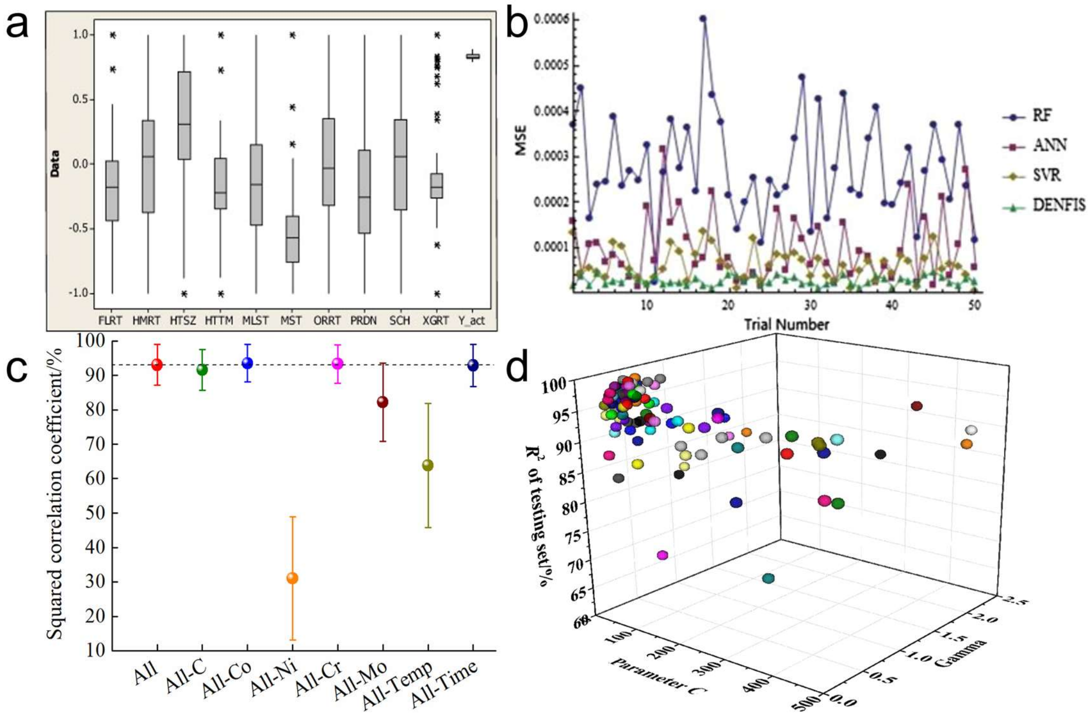


Fig. 4 (a) Box plots of various input-output parameters in the steelmaking process. (b) MSE profiles for different trials during training. Reproduced with permission from Ref. [42]. Copyright 2015, Elsevier. (c) Change in the coefficient of determination when

each feature is removed from the dataset as an input parameter. (d) Test set R^2 values for 100 SVR models trained on different partitions of the training and test sets. Reproduced with permission from Ref. [43]. Copyright 2019, Elsevier.

4.3. Recycling and Resource Recovery

Non-ferrous metals serve as fundamental raw materials for economic development, holding a crucial position in the sustainable growth of the global economy. Given the finite nature of resource reserves, the recovery of waste materials and resource reuse becomes particularly vital. Methods for recovering metals from solid waste primarily include pyrometallurgy, hydrometallurgy, and bioleaching [44]. Pyrometallurgy and hydrometallurgy extract metallic elements from solid waste through chemical reactions such as oxidation, smelting, leaching, and extraction at high temperatures or in aqueous solutions, respectively. Bioleaching, meanwhile, leverages microbial metabolic activity to recover metals. Machine learning techniques have been extensively applied in the recycling of non-ferrous metal resources. In 2024, Bernicky et al. [45] established the ANN model shown in **Fig. 5(a)**, capable of identifying minerals with mass percentages exceeding 2.5%. Finally, the heatmap of the correlation coefficient R^2 in **Fig. 5(b)** shows that the correlation coefficients between temperature and other descriptors are less than 0.08. This indicates that the model effectively decouples temperature signals from molecular and atomic feature signals, ensuring the reliability of subsequent descriptor analysis and providing valuable guidance for pyrometallurgical process parameters. In 2025, Shanshan et al. [46] employed a Random Forest Regression (RFR) model to predict the leaching rates of non-ferrous metals—including lithium (Li), cobalt (Co), manganese (Mn), and nickel (Ni)—from waste lithium batteries in a roasting-water leaching process. They achieved correlation coefficients exceeding 0.8 while establishing a graphical user interface (GUI). In 2022, Mokarian et al. [47]investigated metal recovery in bioleaching processes. By comparing different models, they found that the regression decision tree models demonstrated excellent predictive performance for metal leaching rates, achieving an accuracy of 77%. This demonstrates the significant guiding role of machine learning technology in various non-ferrous metal recovery methods. In the foreseeable future, with the continuous advancement of machine learning technology, the resource utilization rate and economic benefits of non-ferrous metals will undoubtedly see substantial improvement. Concurrently, the integration of approaches such as the Sherwood principle [48], deep learning methods [49], and energy-dispersive X-ray fluorescence spectroscopy (EDXRF) [50]-illustrated in **Fig. 5(c)** and **Fig. 5(d)** has substantially enhanced model predictive capabilities. These advancements have made an indelible contribution to alleviating global supply pressures for fundamental raw materials.

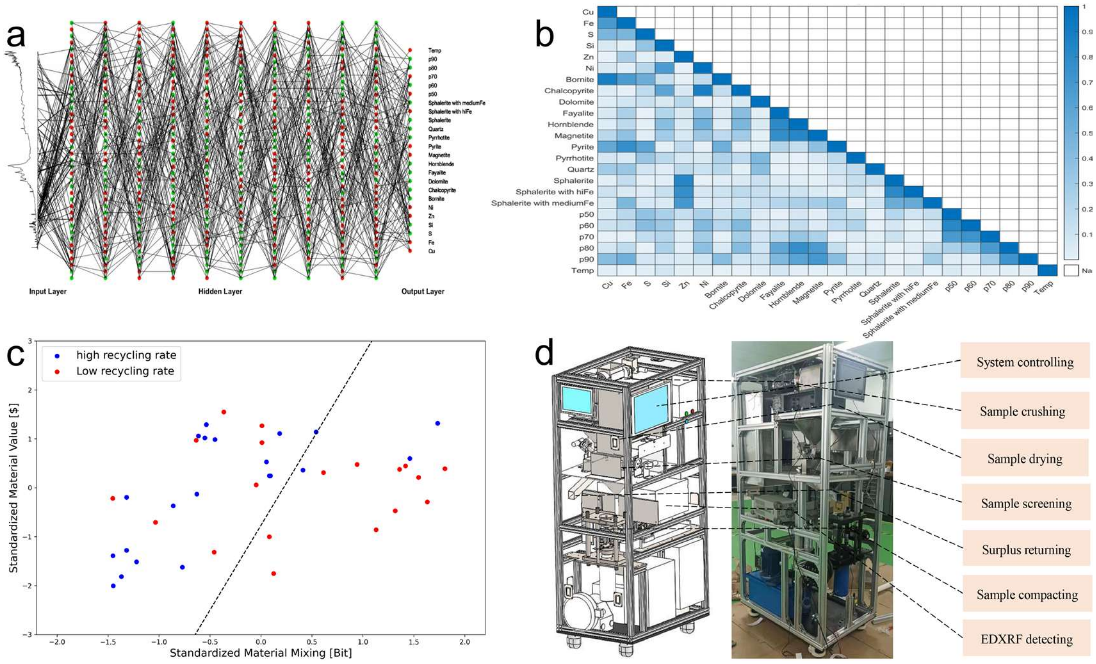


Fig. 5 (a) Optimal combinatorial ANN generated by the training algorithm, depicted by connections with the strongest magnitude. (b) Descriptor correlation R^2 for all descriptors in the combinatorial ANN. Reproduced with permission from Ref. [45]. Copyright 2025, The Royal Society of Chemistry. (c) Sherwood plot of the extended dataset. Reproduced with permission from Ref. [48]. Copyright 2021, Elsevier. (d) Schematic diagram of an online EDXRF detection device. Reproduced with permission from Ref. [50]. Copyright 2023, Elsevier.

4.4. Pollution Control and Environmental Remediation

The primary sources of non-ferrous metal pollution include four major activities: mining, smelting, manufacturing, and recycling. The resulting pollution severely damages soil ecosystems and poses significant risks to human health, with heavy metal contamination being particularly hazardous. Consequently, pollution control efforts primarily focus on heavy metal elements. According to WHO reports, approximately 1.5 million people died from lead poisoning in 2021, primarily due to cardiovascular effects, underscoring the urgency of pollution monitoring and remediation. To date, various soil remediation technologies have been developed, including in-situ remediation techniques such as surface capping, electrokinetic extraction, and bioremediation; and ex-situ remediation methods such as landfill, solidification, and vitrification [51]. These involve a series of physical, chemical, and biological processes to remediate metal contamination through methods such as immobilization, adsorption, and extraction. Due to its ability to identify patterns in large, multidimensional complex data, ML is widely applied in monitoring heavy metal pollution in soil. For instance, in 2023, Zhao et al. [52] used RF, ANN, and XGBoost models to predict the spatial distribution of heavy metals like Cu and Cd. The distribution of Cd is shown in **Fig. 6(a)**; this study successfully identified the significance of anthropogenic factors—such as land use types, road density, population density, and industrial activities—alongside natural factors such as topographic attributes and climatic conditions in influencing heavy metal distribution in soil. Machine learning's unique advantage also enables analysis and simulation of heavy metal migration pathways. For instance, in Wang et al.'s study [53], heavy metal similarity calculations and random walk algorithms were integrated with machine learning. This hybrid framework successfully revealed the migration pathways of As, Cd, Hg, and Pb in soil, with the Cd migration pathway depicted in **Fig. 6(b)**. Collectively, these research findings provide crucial guidance for monitoring metal pollution in soil. Concurrently, various machine learning algorithms capable of uncovering underlying patterns are extensively applied in heavy metal remediation. ML techniques are continuously being leveraged for improvements in areas such as remediation efficacy prediction, material synthesis design, and reaction parameter optimization [54]. Sun et al. employed RF and ANN models to predict heavy metal remediation efficacy in biochar-treated soils, with the ANN model yielding the best prediction performance ($R^2 = 0.841$). They identified heavy metal type and soil pH as the most influential features affecting remediation efficacy [55]. Zhang et al. [56] employed the XGBoost algorithm combined with Bayesian optimization to accurately predict the adsorption coefficient of hydrogel adsorbents for toxic metals. They further utilized SHAP importance analysis to determine the optimal reaction parameters for hydrothermally synthesized hydrogels. ML methods can rapidly predict the potential patterns of non-ferrous metal contamination in soils and optimize the selection of remediation strategies, offering efficient and economical solutions for soil heavy metal pollution management. In summary, machine learning technology provides a new breakthrough for tackling challenging soil non-ferrous metal pollution issues and pioneer's innovative pathways for environmental protection and pollution control of non-ferrous metals.

4.5. Materials Development and Application

Beyond addressing pollution, it is equally important to explore how non-ferrous metals can leverage their unique physicochemical properties to support environmental protection. The application of non-ferrous metal materials constitutes an indispensable component of the product lifecycle for non-ferrous metals. Non-ferrous metals are highly valued for their exceptional electrical conductivity, thermal conductivity, ductility, and corrosion resistance. These properties enable their broad application prospects across sectors, including power generation, transportation, chemical engineering, healthcare, aerospace, and electronic information. This study focuses on exploring the application of non-ferrous metal materials in environmental fields, particularly within water treatment technologies. Materials such as metal-organic frameworks (MOFs)[57], intermetallic compound nanocrystals[58], and organically templated metal oxides[59] play crucial roles in adsorption, catalysis, flocculation, and membrane separation techniques. The introduction of machine learning technology has opened new avenues for the efficient development of novel materials. As shown in **Fig. 6(c)**, machine learning can represent molecules using bond pockets, fingerprints, electron density, symmetry functions, and chemical environment [60]. This significantly expands the application scope of machine learning in materials development while laying a solid foundation for subsequent materials design. In 2024, Fang et al. [61] combined DFT with machine learning methods to develop electrocatalysts for efficient oxygen reduction reaction (ORR) and oxygen evolution reaction (OER), ultimately screening four colored metal materials: $RuCoN_6$, $RuIrN_6$, $OsRhN_6$, and $OsCoN_6$, as shown in **Fig. 6(d)**. Usman et al. [62] employed Gaussian Process Regression (GPR) models to predict the performance efficiency of PDA-s-UiO-66-NH-CM membranes, achieving an accuracy rate of 99%. Comparative analysis with GPR, SVM, and DT models demonstrated that the GPR model exhibited a significantly higher R^2 than SVM and DT models. Thus, the integration of machine learning technologies can continuously support researchers in conducting experiments, accelerate material design and discovery, and significantly shorten R&D cycles. Machine learning-driven high-throughput screening simulation methods are emerging as a crucial force for propelling the advancement of non-ferrous metallurgy.

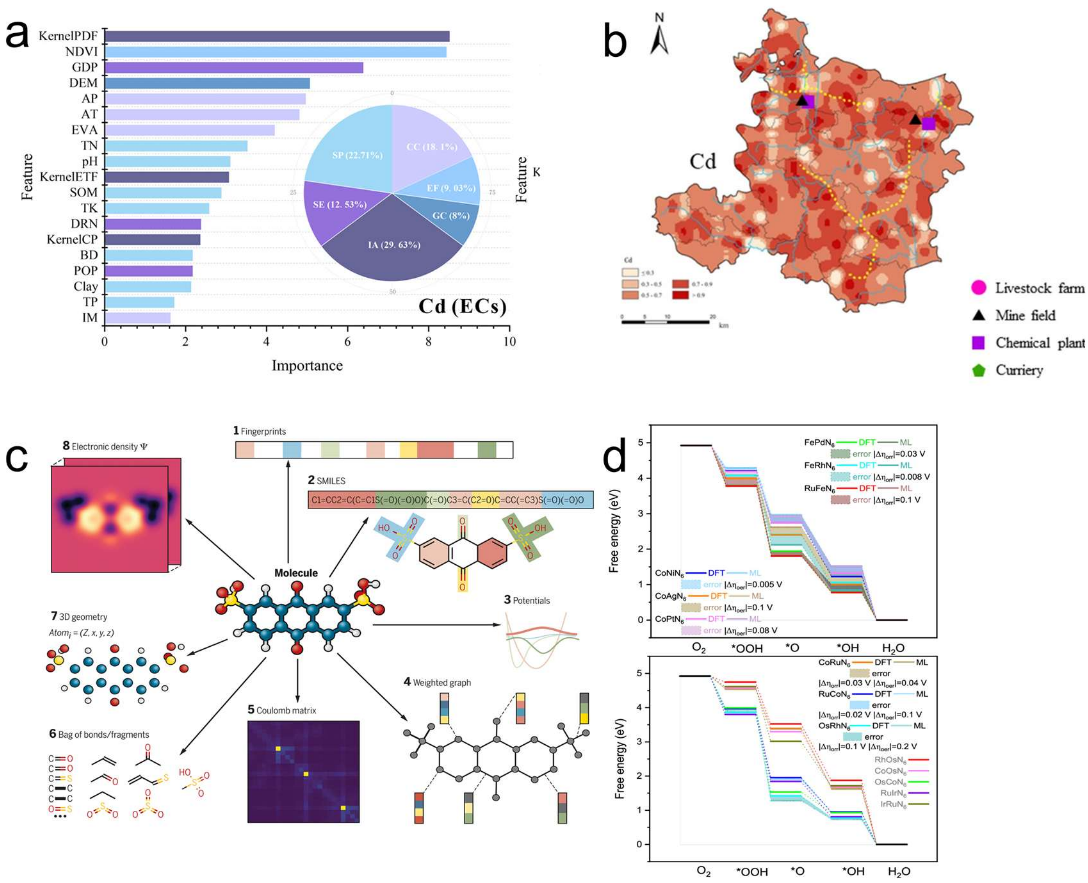


Fig. 6 (a) Results of Cd importance analysis for EC datasets using the RF model. Copyright 2023, American Chemical Society. Reproduced with permission from Ref.

[52]. (b) Soil contamination migration pathways for Cd in the study area. Reproduced with permission from Ref. [53]. Copyright 2023, Elsevier. (c) Molecular representations in machine learning. Reproduced with permission from Ref. [60]. Copyright 2018, American Association for the Advancement of Science. (d) For high-performance ORR or OER catalysts. Reproduced with permission from Ref. [61]. Copyright 2024, American Chemical Society.

In summary, applying machine learning to the entire lifecycle of non-ferrous metals relies on high-quality data, with distinct data issues and algorithmic challenges arising at different processing stages. For mineral exploration and prediction, data primarily originates from remote sensing, geological surveys, and geochemical analysis. Such data is typically noisy, lacks annotated samples, and exhibits significant variability due to differing equipment, collection methods, and data standards across projects. This variability reduces prediction accuracy during large-scale data processing. When applying machine learning to metallurgical processes and optimization, data noise and equipment failures are common issues. Classic regression algorithms struggle to handle the complex linear relationships inherent in metallurgical processes, making deep learning a superior choice. However, the vast amounts of data required for training deep learning models and the intricate relationships between parameters present a major challenge. Compared to metallurgical process optimization, the challenges in non-ferrous metal recycling processes are more focused on mineral identification and classification. Variations in scrap sources, compositions, and states lead to significant data quality disparities. Furthermore, traditional classification algorithms like KNN and SVM perform poorly with complex scrap compositions, hindering machine learning's practical application. In non-ferrous metal pollution remediation, pollutant diversity and standardized monitoring data pose critical challenges. Traditional regression algorithms struggle to accurately capture the dynamic distribution of complex pollutants, while deep learning requires massive datasets—incomplete data often leads to overfitting. For non-ferrous metal material development, the greatest challenge lies in acquiring the experimental data required by machine learning algorithms. The next challenge involves dimensionality reduction of multi-dimensional material data to prevent overfitting. Throughout the entire lifecycle of non-ferrous metals, machine learning applications face data issues and algorithmic bottlenecks. However, with advancements in data collection technologies like remote sensing and smart sensors, coupled with the rise of deep learning techniques such as CNNs, machine learning is assuming an increasingly vital role across the entire lifecycle of non-ferrous metals. Future breakthroughs are anticipated to overcome current limitations and propel industry development.

5. Summary and outlook

This paper systematically reviews the application of ML technology across the entire lifecycle of non-ferrous metals. It highlights the core value of ML in reducing experimental and economic costs while minimizing the reliance on time and luck within the non-ferrous metals industry. The focus is primarily on ML's role in non-ferrous metal exploration, smelting, recycling, pollution control systems, and material applications. To date, ML technology has demonstrated immense potential in the non-ferrous metals sector. By uncovering linear relationships within complex, multi-source data, ML not only enhances the accuracy and efficiency of mineral resource exploration—enabling high-throughput screening and feature importance analysis—but also provides innovative methods and tools for metal smelting, recycling, and pollution monitoring and control. Simultaneously, this technology also faces several shortcomings and challenges. The limitations of applying machine learning throughout the non-ferrous metal’s lifecycle primarily stem from varying operational scenarios, differing data collection methods, and inconsistent data quality due to complex mineral compositions. The entire lifecycle—spanning ore to finished products to waste materials—involves diverse sources with significant compositional variations. Moreover, the high cost and lengthy time required to obtain labeled data from processes like exploration, smelting, and recycling severely constrain the training and validation effectiveness of supervised learning models. The introduction of domain identification methods based on uniform surface approximation and projection for dimensionality reduction and kernel density estimation, alongside linear modeling tools, offers new possibilities for enhancing model generalization and interpretability. Recent advancements in natural language processing may also present new opportunities for data mining and model performance enhancement. Concurrently, the emergence of multiscale modeling and large-scale artificial intelligence will likely become significant research focus in the foreseeable future. Substantial computational resource consumption remains both a cornerstone and a challenge for future AI engineering applications. Moving forward, interdisciplinary collaboration among experts across various fields is essential. Efforts should focus on establishing cross-disciplinary cooperation frameworks and standardized databases to accelerate technology transfer, propelling the non-ferrous metals industry into a new phase of development.

CRediT authorship contribution statement

Haiyang Liao: Manuscript draft, Drawing pictures. **Lei Huang**: Drawing pictures. **Zhenxing Wang**: Design, English revision. **Zhen Zeng**: Design, Manuscript draft, Drawing pictures. **Xue Jia**: Manuscript revision. **Qisheng Huang**: Guidance, Funding acquisition. **Jia Yan**: English summary. **Meng Li**: Data statistics, English summary. **Hongguo Zhang**: Guidance, Editing review. **Zhenxin Chen:** Manuscript revision. **Hao Li:** Guidance, Design.

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

Young Scientific and Technological Talents Training Program of Guangdong Provincial Association for Science and Technology (SKXRC2025249) from Guangdong Provincial Youth Science and Technology Innovation Research Association, Basic and Applied Basic Research Foundation of Guangdong Province (2023A1515010038), National Natural Science Foundation of China(52470070), Guangzhou University Large-scale Equipment Open Sharing Fund, Provincial Natural Science Foundation of Guangdong(2025A1515012030), Youth Foundation of SCIES (PM-zx097-202304-147).

Reference

[1] Liu G, Ren Y, Ma W, et al., "Development process and future trends of chemical refining agents' influence on grain refinement in aluminum alloys," *Journal of Materials Research and Technology*, vol. 29, no. 3 pp. 242-57, 2024.

[2] Zhou C, Peng J, Zheng X, et al., "Investigating the corrosion mechanism of AZ31 alloy bimodal microstructure by the quasi-in-situ EBSD characterization and DFT calculation," *Journal of Materials Research and Technology*, vol. 31, no. 5 pp. 2023-36, 2024.

[3] He X, Yan Z, Liang H, Wang D, "Corrosion fatigue acoustic emission characteristics and evaluation of friction stir welding joints of AZ31 magnesium alloy in 3.5 wt.% NaCl solution," *Journal of Materials Research and Technology*, vol. 25, no. 2 pp. 4582-94, 2023.

[4] Su S, Lu Y, "Densified W Cu composite fabricated via laser additive manufacturing," *International Journal of Refractory Metals and Hard Materials*, vol. 87, no. 4 pp. 105122, 2020.

[5] Geng X, Zhang Y, Han Y, et al., "Two-Dimensional Water-Coupled Metallic MoS_2 with Nanochannels for Ultrafast Supercapacitors," *Nano Letters*, vol. 17, no. 3, pp. 1825-32, 2017.

[6] Hu P, Long M, "Cobalt-catalyzed sulfate radical-based advanced oxidation: A review on heterogeneous catalysts and applications," *Applied Catalysis B: Environmental*, vol. 181, pp. 103-17, 2016.

[7] Liu D, Tian G, Yang Z, et al., "Stress corrosion behavior of 2195-T8 Al-Li alloy with an artificial pit exposed to a 30vol% HNO_3 solution," *Chinese Journal of Aeronautics*, vol. 36, no. 2, pp. 304-15, 2023.

[8] Wang M, Wang L, Pang K, et al., "Understanding stress corrosion cracking behavior of 7085-T7651 aluminum alloy in polluted atmosphere," *Chinese Journal of Aeronautics*, vol. 36, no. 8, pp. 408-21, 2023.

[9] Zhang L, Li H, Bian T, et al., "Advances and challenges on springback control for creep age forming of aluminum alloy," *Chinese Journal of Aeronautics*, vol. 35, no. 10, pp. 8-34, 2022.

[10] Ma S, Zhang Y, Lv J, et al., "Big data driven predictive production planning for energy-intensive manufacturing industries," *Energy*, vol. 211, no. 6, 2020.

[11] Ramprasad R, Batra R, Pilania G, et al., "Machine learning in materials informatics: recent applications and prospects," *npj Computational Materials*, vol. 3, no. 13, pp. 54, 2017.

[12] Zhang S, Xiao K, Carranza E J M, et al., "Integration of auto-encoder network with density-based spatial clustering for geochemical anomaly detection for mineral exploration," *Computers & Geosciences*, vol. 130, pp. 43-56, 2019.

[13] Duan J, Liu X, Wu X, Mao C, "Detection and segmentation of iron ore green pellets in images using lightweight U-net deep learning network," *Neural Computing and Applications*, vol. 32, no. 10, pp. 5775-90, 2019.

[14] Zhang R, Yang J, "State of the art in applications of machine learning in steelmaking process modeling," *International Journal of Minerals, Metallurgy and Materials*, vol. 30, no. 11, pp. 2055-75, 2023.
[15] Xu C, Fu L, Lin T, et al., "Machine learning in petrophysics: Advantages and limitations," *Artificial Intelligence in Geosciences*, vol. 3, pp. 157-61, 2022.
[16] Bomrah S, Uddin M, Upadhyay U, et al., "A scoping review of machine learning for sepsis prediction- feature engineering strategies and model performance: a step towards explainability," *Critical Care*, vol. 28, no. 1, pp. 180, 2024.
[17] Plotnikova V, Dumas M, Milani F, "Adaptations of data mining methodologies: a systematic literature review," *PeerJ Computer Science*, vol. 6, pp. 267, 2020.
[18] Wan X, Li Z, Yu W, et al., "Machine Learning Paves the Way for High Entropy Compounds Exploration: Challenges, Progress, and Outlook," *Advanced Materials*, vol. 37, no. 31, pp. 1521-4095, 2023.
[19] Wong T-T, Yeh P-Y, "Reliable Accuracy Estimates from k-Fold Cross Validation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 8, pp. 1586-94, 2020.
[20] Yates L A, Aandahl Z, Richards S A, Brook B W, "Cross validation for model selection: A review with examples from ecology," *Ecological Monographs*, vol. 93, no. 1, pp. 1557, 2023.
[21] Lu K, Chang D, Lu T, et al., "Machine Learning Model for High-Throughput Screening of Perovskite Manganites with the Highest Néel Temperature," *Journal of Superconductivity and Novel Magnetism*, vol. 34, no. 7, pp. 1961-9, 2021.
[22] Shirmard H, Farahbakhsh E, Müller R D, Chandra R, "A review of machine learning in processing remote sensing data for mineral exploration," *Remote Sensing of Environment*, vol. 268, 2022.
[23] McCoy J T, Auret L, "Machine learning applications in minerals processing: A review," *Minerals Engineering*, vol. 132, pp. 95-109, 2019.
[24] Xiong Y, Zuo R, "GIS-based rare events logistic regression for mineral prospectivity mapping," *Computers & Geosciences*, vol. 111, pp. 18-25, 2018.
[25] Wang C, Chen J, Ouyang Y, "Determination of Predictive Variables in Mineral Prospectivity Mapping Using Supervised and Unsupervised Methods," *Natural Resources Research*, vol. 31, no. 4, pp. 2081-102, 2022.
[26] Wang J, Zuo R, Xiong Y, "Mapping Mineral Prospectivity via Semi-supervised Random Forest," *Natural Resources Research*, vol. 29, no. 1, pp. 189-202, 2019.
[27] Yang F, Wang Z, Zuo R, et al., "Quantification of Uncertainty Associated with Evidence Layers in Mineral Prospectivity Mapping Using Direct Sampling and Convolutional Neural Network," *Natural Resources Research*, vol. 32, no. 1, pp. 79-98, 2022.
[28] Farhadi S, Tatullo S, Konari M B, Afzal P, "Evaluating StackingC and ensemble models for enhanced lithological classification in geological mapping," *Journal of Geochemical Exploration*, vol. 260, 2024.
[29] Daviran M, Ghezelbash R, Hajihosseinlou M, Maghsoudi A, "Uncertainty quantification in genetic algorithm-optimized artificial intelligence-based mineral prospectivity models: automated hyperparameter tuning for support vector machines

and random forest," *Modeling Earth Systems And Environment*, vol. 11, no. 1, pp. 10, 2025.
[30] Mat Isa N A, Mamat W M F W, "Clustered-Hybrid Multilayer Perceptron network for pattern recognition application," *Applied Soft Computing*, vol. 11, no. 1, pp. 1457-66, 2011.
[31] Bhat N, Birbilis N, Barnard A S, "Unsupervised learning and pattern recognition in alloy design," *Digital Discovery*, vol. 3, no. 12, pp. 2396-416, 2024.
[32] Zheng J, Jia R, Liu S, et al., "Sample-efficient reinforcement learning with knowledge-embedded hybrid model for optimal control of mining industry," *Expert Systems with Applications*, vol. 254, 2024.
[33] Tahmasebi P, Javadpour F, Sahimi M, "Data mining and machine learning for identifying sweet spots in shale reservoirs," *Expert Systems with Applications*, vol. 88, pp. 435-47, 2017.
[34] Shafighfard T, Kazemi F, Bagherzadeh F, et al., "Chained machine learning model for predicting load capacity and ductility of steel fiber–reinforced concrete beams," *Computer-Aided Civil and Infrastructure Engineering*, vol. 39, no. 23, pp. 3573-94, 2024.
[35] Rimal Y, Sharma N, Alsadoon A, "The accuracy of machine learning models relies on hyperparameter tuning: student result classification using random forest, randomized search, grid search, bayesian, genetic, and optuna algorithms," *Multimedla Tools And Applications*, vol. 83, no. 30, pp. 74349-64, 2024.
[36] Shao Z, Er M J, "Efficient Leave-One-Out Cross-Validation-based Regularized Extreme Learning Machine," *Neurocomputing*, vol. 194, no. pp. 260-70, 2016.
[37] Qian Wang, Jianmin Bian, Enze Ma "Predicting sorption of organic pollutants on soils with interpretable machine learning," Environ. Pollut., vol. 382, no. 1, pp. 73, 2025.
[38] Maitre J, Bouchard K, Bédard L P, "Mineral grains recognition using computer vision and machine learning," *Computers & Geosciences*, vol. 130, pp. 84-93, 2019.
[39] Sun G, Zeng Q, Zhou J-X, "Machine learning coupled with mineral geochemistry reveals the origin of ore deposits," *Ore Geology Reviews*, vol. 142, 2022.
[40] Durodola J F, "Machine learning for design, phase transformation and mechanical properties of alloys," *Progress in Materials Science*, vol. 123, 2022.
[41] Xiong J, Shi S-Q, Zhang T-Y, "Machine learning of phases and mechanical properties in complex concentrated alloys," *Journal of Materials Science & Technology*, vol. 87, pp. 133-42, 2021.
[42] Laha D, Ren Y, Suganthan P N, "Modeling of steelmaking process with effective machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 10, pp. 4687-96, 2015.
[43] Shen C, Wang C, Wei X, et al., "Physical metallurgy-guided machine learning and artificial intelligent design of ultrahigh-strength stainless steel," *Acta Materialia*, vol. 179, pp. 201-14, 2019.
[44] S. Zhao, W. He, G. Li, Recycling of Spent Lithium-Ion Batteries Processing Methods and Environmental Impacts, Hong Kong, China: Springer, 2019.

[45] Bernicky A R, Davis B, Kadiyski M, Loock H P, "Mineralogical Analysis of Solid-Sample Flame Emission Spectra by Machine Learning," *Anal Chem*, vol. 96, no. 49, pp. 19330-8, 2024.
[46] E S, Niu B, Liu J, et al., "Intelligent metal recovery from spent Li-ion batteries: machine learning breaks the barriers of traditional optimizations," *Green Chemistry*, vol. 27, no. 9, pp. 2478-92, 2025.
[47] Mokarian P, Bakhshayeshi I, Taghikhah F, et al., "The advanced design of bioleaching process for metal recovery: A machine learning approach," *Separation and Purification Technology*, vol. 291, 2022.
[48] Deng S, Zhou X, Huang A, et al., "Evaluating economic opportunities for product recycling via the Sherwood principle and machine learning," *Resources, Conservation and Recycling*, vol. 167, 2021.
[49] Koinig G, Kuhn N, Fink T, et al., "Deep learning approaches for classification of copper-containing metal scrap in recycling processes," *Waste Manag*, vol. 190, pp. 520-30, 2024.
[50] Teng J, Shi Y, Liu Z-H, et al., "Rapid identification of hazardous heavy metal-containing waste by combining EDXRF with machine learning: Taking zinc smelting waste as an example," *Resources, Conservation and Recycling*, vol. 198, 2023.
[51] Liu L, Li W, Song W, Guo M, "Remediation techniques for heavy metal-contaminated soils: Principles and applicability," *Sci Total Environ*, vol. 633, pp. 206-19, 2018.
[52] Zhao W, Ma J, Liu Q, et al., "Accurate Prediction of Soil Heavy Metal Pollution Using an Improved Machine Learning Method: A Case Study in the Pearl River Delta, China," *Environ Sci Technol*, vol. 57, no. 46, pp. 17751-61, 2023.
[53] Wang F, Huo L, Li Y, et al., "A hybrid framework for delineating the migration route of soil heavy metal pollution by heavy metal similarity calculation and machine learning method," *Sci Total Environ*, vol. 858, 2023.
[54] Zheng Y, Sun S, Liu J, et al., "Application of machine learning for material prediction and design in the environmental remediation," *Chinese Chemical Letters*, vol. 36, no. 9, pp. 1238, 2024.
[55] Sun Y, Zhang Y, Lu L, et al., "The application of machine learning methods for prediction of metal immobilization remediation by biochar amendment in soil," *Sci Total Environ*, vol. 829, 2022.
[56] Zhang J, Fu K, Wang D, et al., "Refining hydrogel-based sorbent design for efficient toxic metal removal using machine learning-Bayesian optimization," *J Hazard Mater*, vol. 479, 2024.
[57] Wang J, Liu J, Wang H, et al., "A comprehensive transformer-based approach for high-accuracy gas adsorption predictions in metal-organic frameworks," *Nat. Commun.*, vol. 15, 2024.
[58] Gao Q, Pillai H S, Huang Y, et al., "Breaking adsorption-energy scaling limitations of electrocatalytic nitrate reduction on intermetallic CuPd nanocubes by machine-learned insights," *Nat. Commun.*, vol. 13, pp.1-12, 2022.
[59] Raccuglia P, Elbert K C, Adler P D F, et al., "Machine-learning-assisted materials discovery using failed experiments," *Nature*, vol. 533, no. 7601, pp. 73-6, 2016.

[60] Sanchez-Lengeling B, Aspuru-Guzik A, "Inverse molecular design using machine learning: Generative models for matter engineering," *Science*, vol. 361, no. 6400, pp. 360-5, 2018.
[61] Fang Z, Li S, Zhang Y, et al., "The DFT and Machine Learning Method Accelerated the Discovery of DMSCs with High ORR and OER Catalytic Activities," *The Journal of Physical Chemistry Letters*, vol. 15, no. 1, pp. 281-9, 2024.
[62] Usman J, Abba S I, Baig N, et al., "Design and Machine Learning Prediction of In Situ Grown PDA-Stabilized MOF (UiO-66-NH_2) Membrane for Low-Pressure Separation of Emulsified Oily Wastewater," *ACS Applied Materials & Interfaces*, vol. 16, no. 13, pp. 16271-89, 2024.