Green Chemistry

PAPER

Check for updates

Cite this: *Green Chem.*, 2024, **26**, 7857

Machine learning models accelerate deep eutectic solvent discovery for the recycling of lithium-ion battery cathodes†

Fengyi Zhou,‡[a] Dingyi Shi,‡[b] Wenbo Mu,[c] Shao Wang,[a] Zeyu Wang,[a] Chenyang Wei,[a] Ruiqi Li*[b] and Tiancheng Mu *[a]

Deep eutectic solvents (DESs) have been widely applied to recover spent lithium-ion batteries (LIBs); however, developing effective and efficient systems for cathode leaching *via* the traditional trial-and-error method requires substantial efforts. This work aims to accelerate the discovery of novel promising DESs by leveraging the conditional Generative Adversarial Network (CGAN). Three databases were constructed: (i) DESs leaching cathodes, (ii) DESs leaching metal oxides, and (iii) DES properties. The absolute Spearman's rank correlation and agglomerative hierarchical clustering analysis ensured the selection of an optimal feature set for building predictive models. An XGBoost model was developed, achieving remarkable performance (R^2 = 0.9702, MSE = 0.0007) in predicting cathode solubility in DESs. We employed the Shapley additive explanation (SHAP) method to quantify the importance of acidity, coordination, and reducibility of DESs and provide insights into further research. To accelerate time-consuming investigational procedures, a CGAN model was established, rapidly identifying promising DESs like ChCl : Glycolic acid, with excellent agreement between predictions and experimental results. This study offers a general data analysis framework for other metal oxides (*e.g.*, Cu_xO, Fe_xO_y, ZnO) leaching using DESs, enabling accurate solubility prediction and deepening the understanding of cathode leaching mechanisms. The CGAN model significantly accelerates the development of a DES-based process for lithium-ion cathode recycling, saving development time and effort. Overall, this work facilitates the efficient discovery and development of effective DESs for the recovery of valuable metals from spent LIB cathodes.

Received 22nd March 2024,
Accepted 30th May 2024

DOI: 10.1039/d4gc01418a

rsc.li/greenchem

Introduction

In the context of global decarbonization, the rapidly growing renewable energy market has led to unprecedented challenges for the supply of lithium-ion batteries (LIBs), particularly with regard to raw materials like nickel, cobalt, and lithium.[1,2] Metal supplies are affected by social and geographical factors; for instance, cobalt is available primarily as a companion metal from a limited number of geopolitical concentrated ore deposits,[3] rendering its supply incapable of responding to rapid changes in demand. By 2040, the future demand for lithium, cobalt, and nickel for lithium-ion batteries in electric vehicles will exceed the current production capacity of these raw materials.[4] Additionally, the accumulation of LIB waste can cause environmental pollution and damage public health,[5] and the indiscriminate exploitation of natural resources is unsustainable. As a result, there is an urgent need to effectively recycle and reuse spent LIBs to protect the environment, reduce pressure on natural resources, and increase economic benefits.[6,7]

Developing an efficient, economically sustainable leaching process is key to recovering strategic metals from LIB materials. To date, pyrometallurgical, hydrometallurgical, and direct recycling are the most popular ways of recycling spent LIBs. The direct recycling process could restore the performance of the spent cathode materials in a non-destructive manner,[8–11] and how to scale up and ensure the quality of each batch of direct recycling products still needs further research. Apart from the commonly used traditional acids in the hydrometallurgy approach, such as HCl, HNO_3, and citric acid, deep eutectic solvents (DESs) have recently emerged as efficient and sustainable leaching solvents.[12,13] A DES is a

[a]*Department of Chemistry, Renmin University of China, Beijing 100872, China. E-mail: tcmu@ruc.edu.cn*
[b]*College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China. E-mail: lir@buct.edu.cn*
[c]*Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA, 92093-0404, USA. E-mail: wmu@ucsd.edu*
†Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4gc01418a
‡These authors contributed equally to this work.

eutectic mixture of a hydrogen bond acceptor (HBA) and hydrogen bond donor (HBD) *via* hydrogen bonding[14,15] which has good thermal and chemical stability and favorable features like eco-friendliness, low cost, and composition adjustability. Recently, several research studies have verified that DESs have reusability and high metal selectivity,[16,17] could extract metals without extra reductants and have similar leaching efficiency to traditional acids. The designability of DESs presents plenty of opportunities for researchers to select and optimize components for spent cathode recovery. Some choline chloride-based DESs have been employed to leach lithium cobalt oxide (LCO) and lithium nickel manganese cobalt oxide (NCM), and the selection of HBDs changes with the demand for acidity and reducibility;[18,19] however, the studied DESs for cathode leaching are just the tip of the iceberg. The traditional development for metal recovery from spent LIB cathodes in DESs is described in Fig. 1a; the trial and error-based approach blocks the path to rapidly discover promising leaching DESs.

Machine learning (ML) models can make accurate predictions and reveal key insights in data mining projects. It is a powerful tool to accelerate molecule design and functional materials discovery and processing for applications such as catalysts,[20] pharmaceutical synthesis,[21] and pretreatment of Li-ion batteries.[22,23] ML approaches also have great performance in DESs, such as property prediction[24] and gas absorption,[25] which could unlock new opportunities for the recovery of spent LIBs through DESs. Great attention has been paid to leaching cathodes through DESs;[12] currently, the mainstream view assumes that low viscosity, high acidity, strong coordination, and reducibility of DESs might be beneficial for efficient leaching,[26] while little work has validated these hypotheses, and we are still lacking quantification of the importance of each property. Therefore, to accelerate the screening of effective DESs and deepen our understanding of the leaching mechanism, it is meaningful to apply this powerful tool in the DES formulation development process.

The tasks of this work were to develop ML models to accurately predict the solubility of $LiCoO_2$ and accelerate the discovery of efficient and sustainable DESs (Fig. 1b). First, three databases were constructed from published works, including the corresponding parameters of cathode leaching *via* DESs, metal oxide leaching *via* DESs, and physicochemical properties

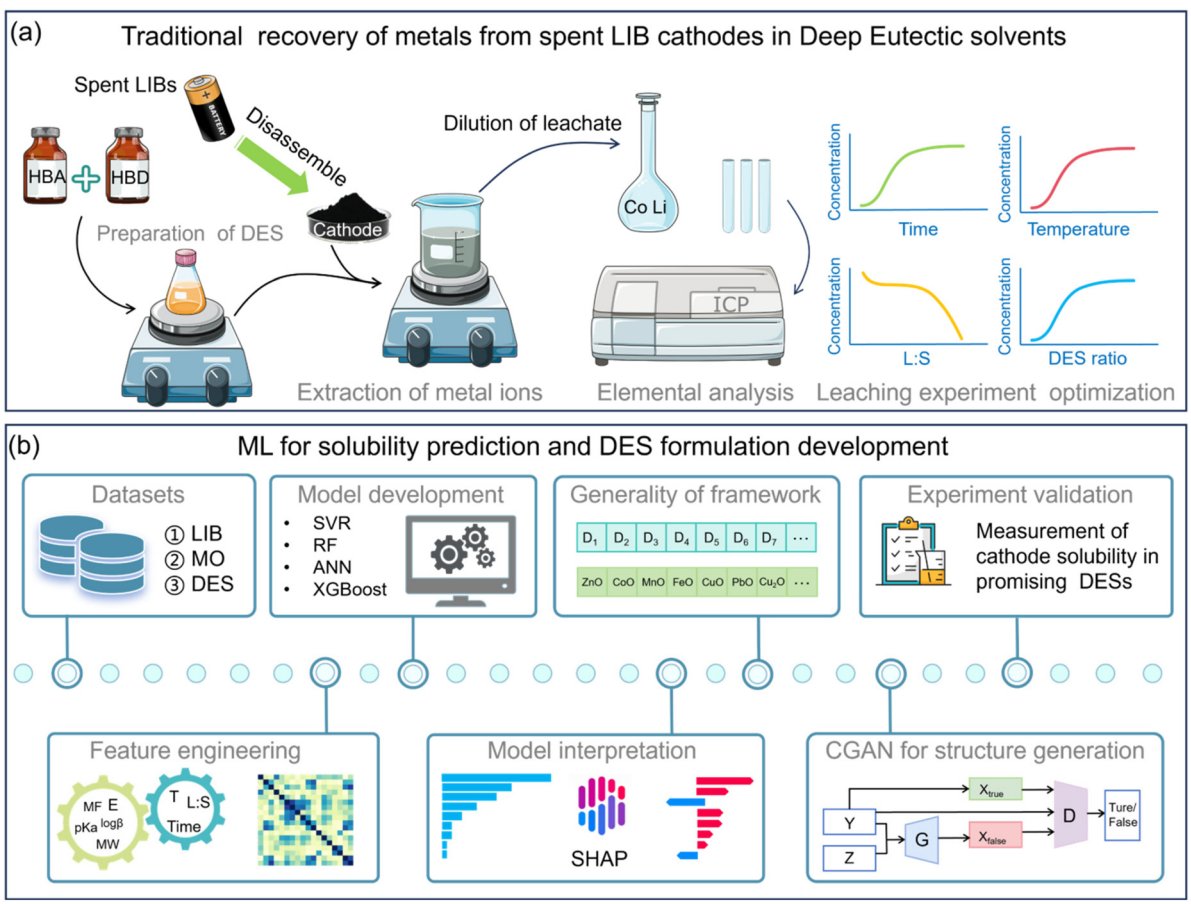

Fig. 1 Schematic demonstrating traditional and ML formulation development approaches for solvometallurgical leaching of lithium-ion battery cathodes using deep-eutectic solvents. (a) Traditional recovery of metals from lithium cobalt oxides in deep eutectic solvents. (b) Workflow shows the process of this study to train and analyze machine learning (ML) models to accelerate the development of DES formulations. CGAN, conditional Generative Adversarial Network.

This journal is © The Royal Society of Chemistry 2024

of DESs, named "LIB" "MO" "DES", respectively. Then, we optimized and screened features by applying absolute Spearman's rank correlation and agglomerative hierarchical clustering analysis for ML models. By training and evaluating four kinds of ML algorithms, including support vector regressor (SVR), random forest (RF), artificial neural networks (ANNs), and extreme gradient boosting (XGBoost), we found that the XGBoost model has the best performance in predicting the solubility of cathodes in various DESs. Thereafter, Shapley additive explanations (SHAPs) were employed to illustrate the XGBoost model and facilitate our understanding of the leaching mechanism. In addition, to verify the generality of the XGBoost model and the selected features, we applied this approach to the MO database to predict the solubility of other metal oxides (*e.g.*, Cu_xO, Fe_xO_y, ZnO) in DESs. Furthermore, based on the SHAP results and DES database, we developed a structure generation model based on CGAN to filter the promising DESs, and the results were verified by experiments.

Experimental

Database

This work established three databases: the database of DESs dissolving lithium-ion battery cathodes termed "LIB", the database of DESs dissolving metal oxides termed "MO data", and the DES database with 36 kinds of DESs are collected for the CGAN model to predict promising systems for dissolving cathodes. The constituent ratio, water content, and viscosity of each DES were collected from previously published studies, and the temperature, time, liquid–solid ratio, and solubility were extracted from figures of leaching efficiency profiles by using the "GetData graph digitizer" application.

All the electrochemical experiments were performed using a CHI 660E electrochemical analyzer. A three-electrode system was implemented at 50 °C: the Ag wire was immersed in 0.01 mol L^{-1} $AgNO_3$ of CH_3CN solution, which acted as the Ag/Ag^+ reference electrode; Pt wire as the counter electrode; ITO-coated glass as the working electrode. Typically, 1 mg $LiCoO_2$ was suspended in 40 μL ethanol, 40 μL deionized water, and 20 μL Nafion solution (5 wt%) to form homogeneous ink assisted by ultrasound. Then 0.1 mL ink was spread onto the surface of ITO-coated glass (Fig. S1†). The electrochemical couple Fc^+/Fc (ferrocenium/ferrocene) was used as an internal reference to calibrate the measured potential. The open circuit potential of each DES was measured at 50 °C and the test lasted for 60 s after 5 min for the system to become stable. The electrochemical data of DESs are summarized in Tables S1 and S2†.

To incorporate as much available data as possible, we used the Web of Science search engine and Google Scholar to comprehensively collect data from previous works. The LIB database contained 19 kinds of DES combinations with 592 values of cathode solubility (Table S3†), and each research article selected for dataset construction includes sufficient data for input features for this work. As shown in Fig. S2,† the data cover a wide range of temperatures (25–200 °C) and the ratio of liquid to solid (2.39–300 g g^{-1}). The MO database contained 11 kinds of DES combinations with 199 values of metal oxide solubility (Table S4†).

Normalization of data can eliminate the influence of dimensions, accelerate model convergence, improve model performance and generalizability, reduce numerical computation errors, and enhance model interpretability. First, we conducted min-max normalization on solubility, with the following formula:

$$S_{\text{normalized}} = \frac{S - S_{\min}}{S_{\max} - S_{\min}}, \quad (1)$$

where S represents the value of each solubility; $S_{\min}$ and $S_{\max}$ indicate the minimum and maximum solubility values, respectively. This method linearly maps the data to an interval. Next, we shuffled the database and divided it into training and testing sets in an 8 : 2 ratio.

Algorithm

Support vector regression (SVR) is a supervised learning algorithm based on statistical learning theory, effective in handling high-dimensional, non-linear, and small-sample data regression problems. Different from SVR, the tree models, *e.g.*, RF and XGBoost provide rich nonlinear variations and can be applied to many complex data sets. For tabular data, Grinsztajn *et al.*[27] demonstrate that tree models have better performance than neural networks. Random Forest (RF) is an ensemble learning algorithm that constructs multiple decision trees and each tree will come up with its own result and finally the results will be averaged on a majority rule. XGBoost is a tree-based ensemble learning algorithm that performs exceptionally well in structured data classification and regression problems. Artificial Neural Network (ANN) is a computational model that mimics the structure and function of a biological neural network, excelling at handling complex non-linear relationships.

Data analysis framework

We developed a unified data analysis framework to predict the solubility of DESs for cobalt ions and to analyze how related features affect solubility. The structure of a substance affects its properties, and different encoding of the structure often leads to different results in the final property prediction. Morgan fingerprint is one of the molecular representations commonly used in cheminformatics and computational chemistry.[28] In our framework, with a radius of 2 and a fixed length of 100, we used the RDKit scientific computing package to calculate the Morgan fingerprints of DES structures. Since the structure of a DES is usually composed of a HBA and HBD in certain proportions, we summed the molecular fingerprints of HBA and HBD in proportion to obtain the coding vector of the overall structure. Based on the chemical knowledge, we selected some features that are more relevant to the solubility of cobalt ions. By calculating Spearman's correlation coefficient between the features and using the agglomerative hier-

archical clustering method[29] to cluster the features, the features with a higher correlation were eliminated. In the process of agglomerative hierarchical clustering, each data point is initially considered as a separate cluster, and then the most similar clusters are merged based on the Euclidean distance until all data points are merged into one cluster. By comparing the predictive effect of the features selected after each layer of clustering to predict solubility, we determined the final features to be used.

After filtering the desired features, the feature groups and structural codes are spliced and input into four models, XGBoost, RF, SVR, and ANN, to predict the value of solubility, and the prediction effects of the 4 models are compared to select the optimal model. During the training process of the model, to fully utilize the data and reduce the risk of overfitting the model, allowing the model to fully learn the regularities behind the data, we used a 10-fold cross-validation technique for model evaluation and learning. This involved dividing the database into 10 mutually exclusive subsets, using 9 subsets for training each time, and then validating the remaining subset. After repeating this process 10 times, the average of the 10 validation results was taken as the final result. 10-fold cross-validation is often combined with parameter tuning to help choose the best parameters for the model, thereby improving the performance and generalization ability of the model. We optimized the hyperparameters of three algorithms SVR, RF, and XGBoost by using the grid search technique, and the grid search range for each algorithm during the model training is shown in Table S5.† SHapley Additive exPlanations (SHAP) is an approach based on the concept of Shapley values from cooperative game theory, used to explain model prediction results. We used the SHAP method to analyze the best model, which could sort the feature importance by SHAP values, reflect the relationship between a feature and the target value, and provide examples to analyze the effects of individual feature values on the system.

Structure generation

As shown in Fig. 2, after obtaining the SHAP analysis results, we set reasonable conditions and used conditional generative adversarial networks (CGAN)[30] to generate ideal Morgan molecular fingerprints for the DES structure. A typical generative adversarial network consists of two basic networks: the generator (G) and the discriminator (D), also known as the discriminative network. However, because the generator is based on random noise, the generated data is often uncontrollable. To ensure that the generated data is controllable, we used conditional generative adversarial networks to generate DES data. The input of the generator is the concatenated set condition and random noise, with the expectation that the generator can generate fake data X with condition y. For the discriminator, whether it discriminates real data X or fake data, the discriminator will combine it with conditional information for discrimination.

Based on the generated molecular fingerprints, we can help filter out DESs with higher solubility in the DES database. The specific method is as follows: firstly, encode the DES in the database into molecular fingerprints, and then, using the generated molecular fingerprints as a reference, calculate the similarity with each fingerprint in the database, where the higher the similarity, the more likely it is to have a higher solubility for cobalt ions.

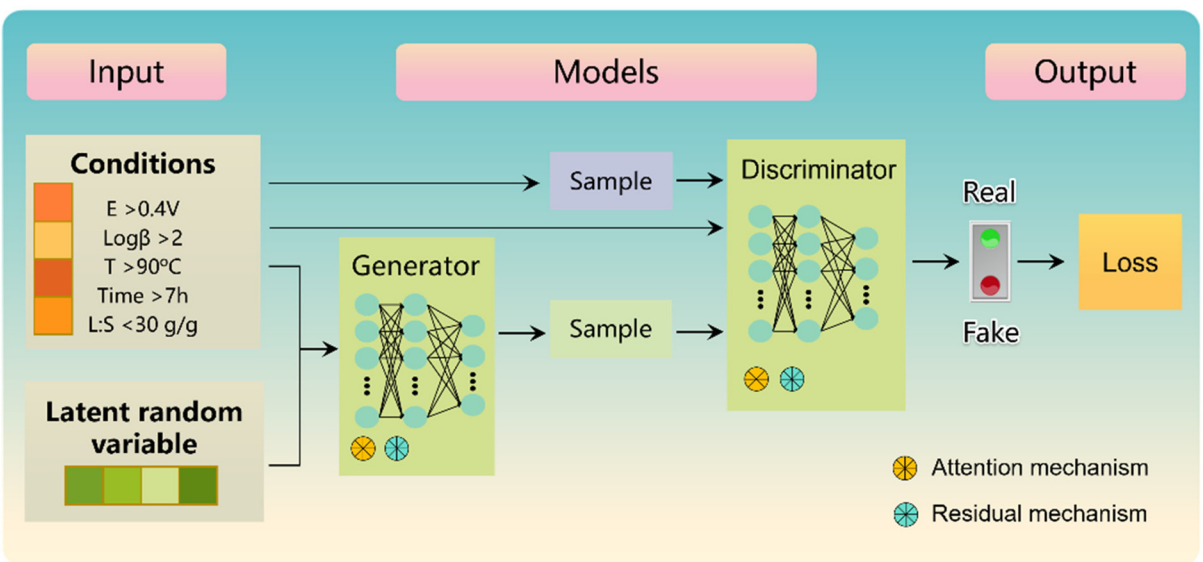

Fig. 2 Conditional generative adversarial networks (CGAN) to generate ideal Morgan molecular fingerprints for DESs. Generators and discriminators have opposite objectives. The generator network takes random noise (usually sampled from a normal distribution) and conditioning information as input, and produces samples in the desired domain. The discriminator network takes samples from both the real data distribution and the generated data distribution (produced by the generator) along with conditioning information and tries to distinguish between them.

This journal is © The Royal Society of Chemistry 2024

Metrics

For the prediction part of the solubility of DES, MSE, and *R*-squared are used to evaluate the prediction effect of four models: ANN, SVR, RF, and XGBoost. MSE is an indicator used to measure the difference between the predicted results of the regression model and the true values. It calculates the average of the squares of the differences between the predicted values and the true values and is one of the most commonly used indicators in regression model evaluation. The formula for calculating MSE is as follows:

$$\mathrm{MSE} = \frac{\sum_{i=1}^{N} (Y_i - \tilde{Y}_i)}{N}, \quad (2)$$

where N represents the sample size, Y_i is the true target value, and $\tilde{Y}_\mathrm{i}$ is the predicted target value. The smaller the MSE value, the smaller the difference between the model's predicted results and the true values, indicating a better predictive performance of the model. *R*-Squared (R^2), also known as the coefficient of determination, is a statistical measure used to assess the fit goodness of a regression model. R^2 represents the proportion of the variability of the dependent variable that can be explained by the model, and its calculation formula is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (Y_i - \tilde{Y}_i)^2}{\sum_{i=1}^{N} (\tilde{Y}_\mathrm{i} - \bar{Y}_i)^2}, \quad (3)$$

where $\bar{Y}_i = \sum_{i=1}^{N} \tilde{}\mathrm{Y_i}/$, represents the average of the model's predicted target values, the R^2 value ranges from 0 to 1, with a value closer to 1 indicating a stronger explanatory power of the model for the dependent variable, *i.e.*, better model fit; while a value closer to 0 indicates poorer model fit, unable to explain the variation in the dependent variable. By combining the two evaluation metrics of MSE and R^2, a more accurate evaluation of the model's predictive ability can be obtained.

As for the structure generation part, since the discriminator's result has only two values, 0 representing fake and 1 representing real, its essence is a binary classification machine learning task. BCEloss is commonly used in machine learning to measure binary classification tasks,[31] typically used to measure the difference between the probability distribution of the output of a binary classification model and the true labels. Its calculation formula is as follows:

$$\mathrm{BCE\,loss} = -\frac{1}{N}\sum_{i=1}^{N}(s_i \log(\tilde{s}_i) + (1 - s_i)\log(1 - \tilde{s}_i)), \quad (4)$$

where N represents the sample size, s_i represents the actual label (0 or 1), and $\tilde{s}_i$ represents the predicted probability of the discriminator. BCELoss measures the difference between the model's output probability and the actual label. When s_i is 1, the BCELoss is $-\log(\tilde{s}_i)$; when s_i is 0, the BCELoss is $-\log(1 - \tilde{s}_i)$. Thus, BCELoss aims to minimize the cross-entropy between the model's prediction and the actual label, thereby improving the model's predictive accuracy for binary classification problems.

Molecular fingerprints are typically high sparse binary vectors with most of their elements being 0. Cosine similarity effectively handles this sparsity by focusing only on the non-zero dimensions of the vectors, making it well-suited for molecular data.[32] During the filtering of a DES, we used cosine similarity to measure the similarity between the generated molecular fingerprints and those in the database, as shown in the following formula:

$$\mathrm{Similarity}\,(F_1 \cdot F_2) = \frac{F_1 \cdot F_2}{||F_1||||F_2||}, \quad (5)$$

where F_1 and F_2 represent two molecular fingerprint vector compositions, $F_1 \cdot F_2$ represents the inner product of these two vectors, $||F_1||$ and $||F_2||$ denote the L2 norms of F_1 and F_2, respectively. The cosine similarity ranges from -1 to 1, measuring the similarity between two vectors by calculating the cosine value of the angle between them. When the angle between two vectors is 90 degrees, the cosine similarity is 0, indicating no similarity. When the angle is 0 degrees, the cosine similarity is 1, indicating identical direction.

Materials

Choline chloride (ChCl, 98%) was acquired from J&K Scientific Ltd. Glycolic acid (98%), and *o*-cresol (98%) were purchased from Aladdin Biochemical Technology Co., Ltd (Shanghai, China). Levulinic acid, glutaric acid (99%), and trimethylamine hydrochloride (TA HCl, 98%) were purchased from Macklin Biochemical Technology Co., Ltd (Shanghai, China). All chemicals were used as received without further purification.

Preparation of DESs

All DESs were prepared by simply mixing components at fixed ratios and heating at 80 °C until a homogeneous, transparent phase had formed.

Metal extraction experiments

Typically, DESs with a certain mass (*e.g.*, 3 g) and 0.1 g $LiCoO_2$ were added into a glass bottle followed by heating isothermally at a specific temperature (*e.g.*, 90 °C) and magnetically stirred. After the dissolution within a specific time, the leaching solutions were taken for centrifugation at 10 000 rpm for 20 min. The liquid phase was used for the measurement of Co concentration. The Co concentration in the liquid phase of the DESs/$LiCoO_2$ mixture was conducted using an inductively coupled plasma optical emission spectrometer (ICP-OES, Agilent Technology Co., Ltd).

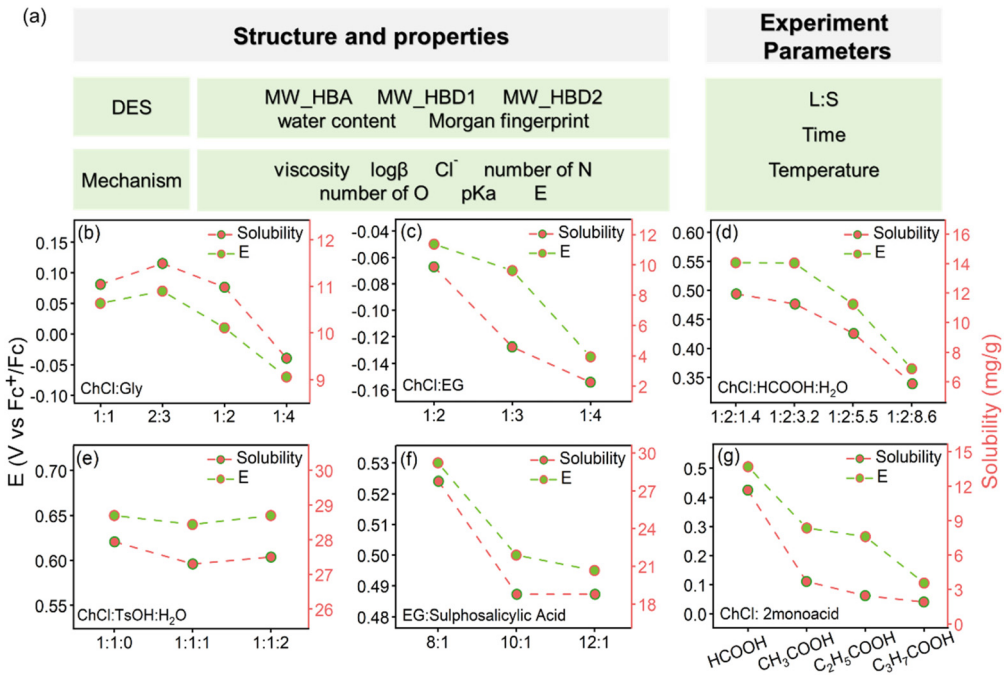

Fig. 3 (a) Feature selection. Structure and properties of DESs, and experiment parameters of the cathode leaching process were selected as descriptors for models to provide information. MW, molecular weight; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; log β: complex stability constant; L : S, the mass ratio of DES to cathode; number of *N*, number of nitrogen atoms; number of *O*, number of oxygen atoms; *E*, potentials of reduction reactions. (b–g) Relationship of potential and solubility of various DESs. The potentials were measured by using LCO as the working electrode. The solubility of cobalt ions in various DESs is collected and converted from previous studies. (b) ChCl : Gly (glycerol),[33] (c) ChCl : EG (ethylene glycol),[17] (d) ChCl : HCOOH : H_2O,[41] (e) ChCl : TsOH : H_2O,[19] (f) EG : sulphosalicylic acid,[35] (g) ChCl : 2 monoacid.[16]

Results and discussion

Feature engineering

An ideal feature set is constructed to offer specific information and help build a model with good performance. As shown in Fig. 3(a), a total of 14 accessible properties and parameters were initially selected as features to predict the solubility of cobalt ions. Simple features such as molecular weight, the water content of the DES, the viscosity of the DES, the ratio of HBA and HBD, and experiment parameters were directly gained from previous work.[16–19,33–46] The Morgan fingerprint was employed to describe the chemical structure,[28] which could be conducive to analysis and interpretation. Besides the physicochemical properties of DESs, the leaching mechanism is critical for the solubility prediction. The mechanism of the process has been analyzed in previously published works, and the key characteristics could be summarized as acidity, viscosity, reducibility, and coordination. The acidity of DES mainly depends on HBD, thus the pK_a of HBD was selected as one of the input features.[47] The coordination of Cl^- and organic acid ions with cobalt ions could facilitate the dissolution of cathodes, and we selected the complex stability constant, the number of coordination atoms oxygen and nitrogen, together with the existence of Cl^- as input features to provide coordination information.

In the leaching process of cathodes, such as LCO, the reduction of Co^{3+} to Co^{2+}, plays a key role in improving the leaching efficiencies of spent LIB cathodes.[12,38] To quantify the reduction reaction and provide information for the model, the reduction reaction potentials were measured and we served it as an electrochemistry descriptor E for the model. The measured reaction potential values of LCO dissolved in different DESs are summarized in Tables S1and S2,† and we displayed the relationship between reaction potential and solubility of Co in Fig. 3b. Under ideal conditions, the reaction could be described as eqn (6), and the Nernst equation is expressed as eqn (7).

$$CoO_2^- + e^- + 4H^+ = Co^{2+} + 2H_2O \quad (6)$$

$$E(CoO_2^-/Co^{2+}) = E^{\theta}(CoO_2^-/Co^{2+}) + \frac{0.0591}{1}\lg\frac{[H^+]^4}{[Co^{2+}]} \quad (7)$$

In reality, the reaction potentials were not only affected by the reducibility and the acidity of the systems, but also the coordination state of Co^{2+} changes with the ligands provided by different DESs. As shown by Fig. 3b–g, reaction potentials of the same components DESs change with different ratios, which suggests the effect of acidity and ligands. In the ChCl : glycerol and ChCl : EG systems (Fig. 3b and c), the acidity of alcohol DESs is stable thus the slight difference in the potentials could be attributed to the changes in ligand content. The effect of changes in acidity can be noticed in Fig. 3d–g, the reaction potentials decrease with the decreasing of acidity when pK_a of HBDs decrease or diluted with water. In

 This journal is © The Royal Society of Chemistry 2024

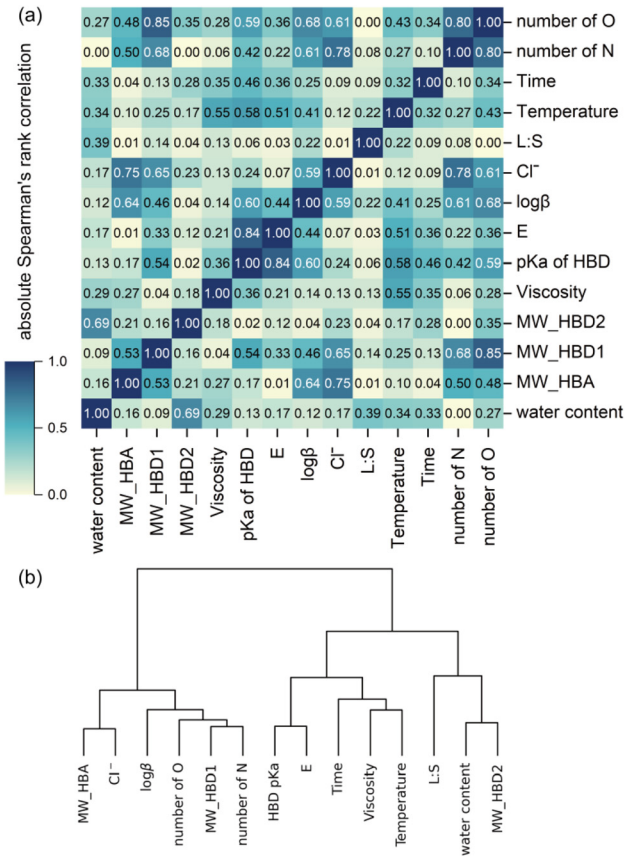

Fig. 4 (a) Heat map of absolute Spearman's rank correlation between the 14 features initially selected. Dark blue indicates a Spearman rank correlation of 1, while light yellow indicates a Spearman rank correlation of 0. (b) A dendrogram obtained through agglomerative hierarchical clustering, representing the hierarchical clustering of features obtained through cluster analysis.

particular, when the acidity of a DES is high enough, the water dilution cannot significantly alter the reaction potential. The strong correlation between acidity and E was also verified in the heat map (Fig. 4). In a word, the electrochemistry descriptor E has a strong correlation with the solubility of Co, which could be helpful for the model and the necessity of E in the model is also verified in the part of model interpretation.

Generally, a comprehensive and detailed feature set is crucial for the model to make accurate predictions, while the use of excessive features can cause redundant problems and degrade the generalization ability of the model. It is necessary to assess and optimize the feature set to reduce model complexity, save computational resources, and improve model performance. Here, we analyzed and refined the 14 input features by a Spearman's rank correlation heat map and agglomerative hierarchical clustering analysis (Fig. 4). The color of each square in the heat map shows the correlation of feature pairs, and dark blue means a strong correlation, such as MW_HBD1, number of N, and number of O (>0.8), Cl^- and MW_HBA (>0.75), Cl^- and number of N (>0.78), E and pKa (>0.84), which reflected the redundancy between input features. To identify the potentially redundant input features, hierarchical clustering was performed on these 14 features. Hierarchical clustering is a distance-based clustering method,[29] which could provide a better insight into the similarities and differences in these features. Here, we gradually merge the data points into larger and larger sets of clusters by calculating the sum of squared deviations between them until some stopping condition is satisfied. As shown in the tree diagram (Fig. 4b), features that join together have similarities, and we removed the similar feature clusters step by step. Table 1 summarizes the feature groups after clustering selection.

Model development

The feature sets determine the performance ceiling of ML, and the algorithms render the model to reach the ceiling as high as possible. Based on the results of feature clustering, we input the 5 feature groups (Table 1) into 4 models XGBoost, RF, SVR, and ANN by combining them with vectors composed of structural codes, respectively. The hyperparameters of three algorithms SVR, RF, and XGBoost were optimized by using the grid search, and the grid search range for each algorithm during the model training is shown in Table S5.†

The overall predictive performance of each model for different feature groups on the test set is summarized in Table 2. Overall, $R^2_{\mathrm{XGBoost}} > R^2_{\mathrm{others}}$, $\mathrm{MSE}_{\mathrm{XGBoost}} < \mathrm{MSE}_{\mathrm{others}}$, which indicates that among the 5 feature groups, compared to the other three models, the generalization of the XGBoost model is the best, and its predictive effect on the test set is the best. In addition, it can be found that among the five feature groups, $R^2_{\mathrm{group\ 1}} > R^2_{\mathrm{others}}$ on the test set, $\mathrm{MSE}_{\mathrm{group\ 1}} < \mathrm{MSE}_{\mathrm{others}}$,[48] and the effect of the ANN model demonstrates this too. It indicates that feature group 1 is the most suitable feature group among the 5 feature groups and also the feature group that makes the model perform best. Compared to all features (group 0), feature group 1 removed 5 redundant features: water content, Viscosity, pKa, Cl^-, and number of N. The removal of pK_a of HBD could be easily explained by eqn (7).

Table 1 Hierarchical clustering selected feature groups

Group 0	Group 1	Group 2	Group 3	Group 4
Water content				
MW_HBA	✓	✓	✓	✓
MW_HBD1	✓	✓		
MW_HBD2	✓	✓	✓	
Viscosity				
pK_a				
E	✓	✓	✓	✓
Log β	✓	✓	✓	
Cl^-				
Number of O	✓			
Number of N				
L : S	✓	✓	✓	✓
Temperature	✓			
Time	✓	✓	✓	

Group 0 is the feature group before the optimization of hierarchical clustering, which has all 14 features. The features contained in each feature group are represented by a check sign.

Table 2 Prediction results on the test set for the four model inputs with different feature groups

Feature group	Model	R^2	MSE
0	SVR	0.7129	0.0066
	RF	0.9553	0.0010
	XGBoost	0.9622	0.0009
	ANN	0.9077	0.0021
1	SVR	0.7287	0.0063
	RF	0.9440	0.0013
	XGBoost	**0.9702**	**0.0007**
	ANN	0.9521	0.0011
2	SVR	0.7340	0.0061
	RF	0.9196	0.0019
	XGBoost	0.9359	0.0015
	ANN	0.9160	0.0019
3	SVR	0.7374	0.0061
	RF	0.9075	0.0021
	XGBoost	0.9445	0.0013
	ANN	0.9271	0.0017
4	SVR	0.7774	0.0051
	RF	0.8518	0.0034
	XGBoost	0.8546	0.0033
	ANN	0.8521	0.0034

Group 0 has all 14 features. All data in the table are results obtained through ten-fold cross-validation. The bold markings in the table represent the best results within each layer, and the bolded entries represent the best results in the overall performance.

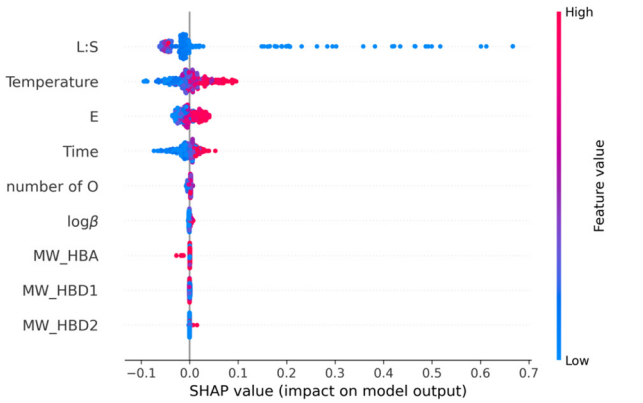

Fig. 5 Shapely additive explanations (SHAP) analysis for the XGBoost model with layer 1 feature group as inputs. The color of each point represents the feature value of a single DES, as shown by the color bar on the right, from dark red to light blue for the relative size of its feature value, respectively. The average of the absolute values of all points can be used as an indicator of feature importance, and the most important feature here is L : S.

While the removal of viscosity could be speculated that the viscosity of DESs is highly correlated with temperature and their structure, thus the effect of viscosity on solubility could be reflected by temperature and Morgan fingerprints,[49,50] and this is consistent with the feature clustering (Fig. 4). Overall, results show that XGBoost with feature group 1 as inputs has the best performance. Specially, with the hierarchical clustering analysis, we minimized the number of features, which not only greatly reduced the computational complexity, but also optimized the predictive performance of models, further demonstrating the strong relationship between the selected features and solubility.

Model interpretation

To quantify the influence of structures and properties of DESs and experimental conditions on the solubility of Co, we utilized the Shapley additive explanation (SHAP) model to explain and analyze the best XGBoost model. With this approach, the global and local interpretability of features could be extracted and thus illustrate the contribution of each feature and guide the discovery of promising DES. Fig. 5 is the scatter plot of SHAP values for the features of group 1, which exhibits the importance ranking of features and the global contribution of each feature based on SHAP values. The SHAP model assigns a SHAP value to each input data point of XGBoost, and the magnitude of its absolute value represents the level of importance for the prediction. In addition, the positive or negative sign represents the positive or negative impact of the feature on the prediction results. It could be observed that the experiment parameters are the most contributing features, such as L : S and temperature. The increasing amount of solids increases the surface area, which is beneficial to promote the leaching reaction.[51] Regarding the properties of DESs, it is confirmed that the measured potential E matters most in the model prediction, and the SHAP value shows that the higher acidity and reducibility (E) and stronger coordination ($\log\beta$) have positive impacts on solubility.

Local interpretability of the XGBoost model is described to observe the specific trends of features. The ability to adapt to varying conditions is essential for ensuring that a model remains relevant and effective across different scenarios. To maximize the model's applicability, it is imperative to not only recognize but actively accommodate these variations.[52] It entails the development of conditions that are not just satisfactory, but optimal, tailored to align with the specific requirements and intricacies of real-world processing tasks. Achieving this level of customization will significantly enhance the model's effectiveness and ensure it delivers accurate, reliable results under diverse operating conditions. We plot the dependency plot of these important features to learn and determine the impact of feature values (Fig. 6a–d). It could be found that the lower values of the ratio of liquid to solid (L : S) have a positive impact on solubility, and the higher temperature and longer leaching time could directly increase solubility. The range of these feature values with positive SHAP values can be extracted from dependency plots, L : S < 30, temperature > 90, time > 7, E > 0.4. Identifying condition ranges with positive effects guides the step of condition optimization to reduce work for leaching experiments. According to the commonly used types of DESs for recovery spent LIBs, type III and type V, two specific instances with high Co solubility are selected to analyze the exact contribution of each feature on the prediction of solubility. A type III DES is the combination of a quaternary ammonium salt and a carboxylic acid, thus we chose *p*-toluenesulfonic acid·H_2O : ChCl, and as shown in Fig. 6e, the

 This journal is © The Royal Society of Chemistry 2024

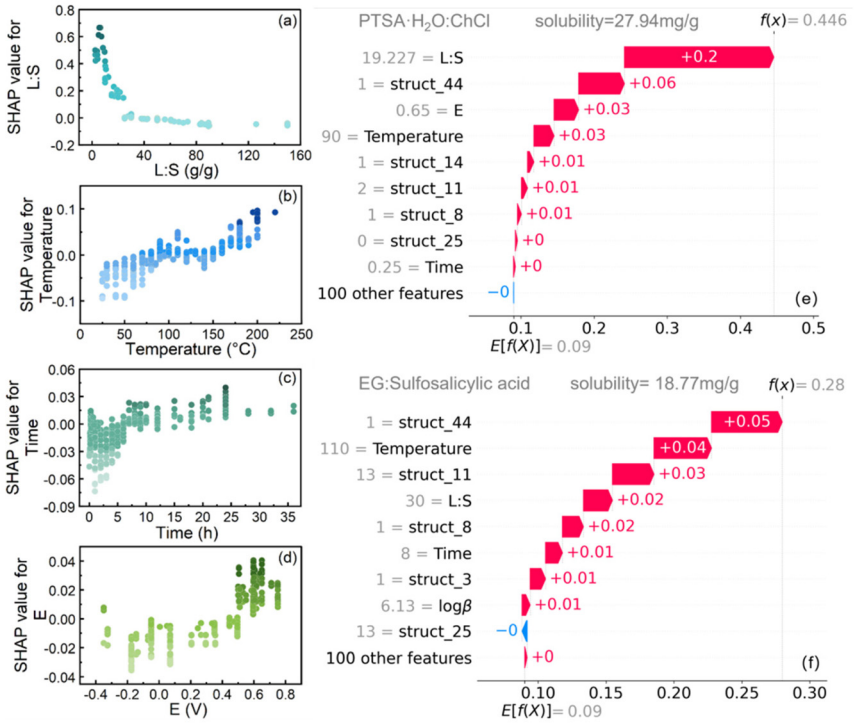

Fig. 6 Local interpretability of the XGBoost model with layer 1 feature group as inputs. The dependence plots of SHAP values with (a) L : S, (b) temperature, (c) time, and (d) *E*. The waterfall plots for selected instances (e) PTSA (*p*-toluenesulfonic acid)·H_2O : ChCl, and (f) EG (ethylene glycol) : sulfosacylic acid.

high *E* and low L : S contribute most in the high solubility prediction. A type V DES is made up of nonionic molecular substances, such as EG:Sulfosalicylic acid (Fig. 6f), the proper experiment conditions contribute positive effects on predictions, which have similar trends. It is worth noting that the Morgan fingerprints play a key role in solubility predictions, for example, both structure 44 and structure 11 (Fig. S3†) have hydroxyl groups, and structure 44 is the favorable structure to coordinated cobalt ions,[35,53] which might underline the importance of coordination. In conclusion, with the SHAP analysis for the XGBoost model, the importance of coordination, acidity, and reducibility is quantified, and rational design for features is verified.

Generality of the model development approach

As discussed above, the data analysis framework we developed has good performance on the LIB database, DESs have abundant functional groups, thus the process of structural encoding could provide comprehensive structural information; the utilization of Spearman correlation coefficient and agglomerative hierarchical clustering could rationally filter the features; SHAP model helps to analyze the contributions of features on predictions and provide guidance for further study. Therefore, we speculate that this data analysis framework is suitable for the machine learning of DESs. To verify the generality of this framework, we further applied it to the MO database which is composed of the dissolution data of metal oxides in DESs (Table S4†). Considering that the layer 1 feature group describes the process of metal leaching in DESs precisely thus it has generality in this field, we used it to describe these common metal oxides such as CuO, ZnO, Fe_2O_3, PbO, and so on. Different from the LIB database, the types of metal oxides are various, thus the lattice energy is added as a descriptor to label different metal oxides. By applying this framework to MO data, we obtained the predictive model and SHAP analysis. It is worth mentioning that the prediction results of XGBoost on the MO database are still optimal compared to the other three models (Table 3). In addition, to show the effect of feature selection on model performance, the prediction accuracy of

Table 3 Predictive performance of four models on the training and testing sets of the MO database

		R^2	MSE
SVR	Train	0.3067	0.0160
	Test	0.0989	0.0370
RF	Train	0.8414	0.0037
	Test	0.7373	0.0108
XGBoost	Train	0.9779	0.0005
	Test	**0.9072**	**0.0038**
ANN	Train	0.4806	0.0120
	Test	0.4432	0.0229

All data in the table are results obtained through ten-fold cross-validation. The bolded entries in the table represent the best results among the four models.

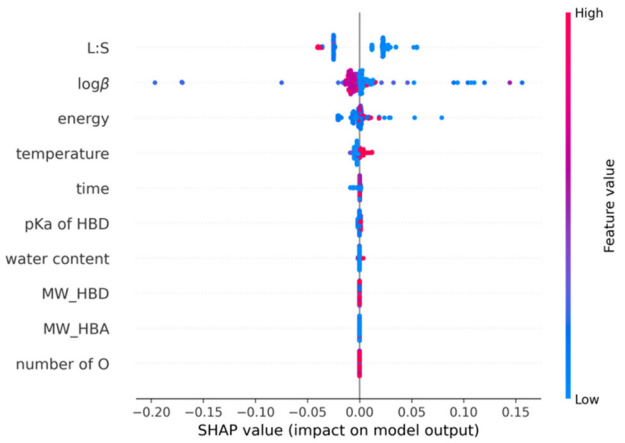

Fig. 7 Density scatter plot of SHAP values for XGBoost model on MO database.

different feature sets on the XGBoost model was compared (Table S7†), and the results show that the optimal feature set has the best performance. Fig. 7 shows the results of the SAHP interpretation of the trained XGBoost model, which verified that selecting lattice energy as a descriptor is necessary, the importance of coordination ability is highlighted again. Hence, it is demonstrated that the optimal feature set and this data analysis framework are robust and fundamental and thus have generality in the field of metal oxides leaching in DESs.

Promising DES prediction and validation

In addition to the guidance for the leaching experiment given by the predictive model and SHAP model, we were interested in rapidly identifying promising DESs for spent LIBs recovery. Thus, the DES database was constructed, comprising information on DESs (Table S2†). This database was applied to a conditional generative adversarial network (CGAN) model. Fig. 2 presents the details of the CGAN model. Unlike the general generative adversarial network, the CGAN applies supervised learning to the generator.[30] To generate fake data X under the specified conditions, the generator's input is a vector composed of random noise and conditional information (Conditions), rather than just random noise. The discriminator combines the data with conditional information for discrimination whether discriminating real data (X_{real}) or fake data (X_{fake}). Based on the results of SHAP analysis and practical considerations, we set Conditions (E = 0.5 V, log β = 3, T = 90 °C, time = 8 h, L : S = 30 g g^{-1}) and employed CGAN to generate the ideal molecular fingerprint under these conditions. The prepared condition vector was concatenated with random noise as the input to the generator, expecting it to generate the theoretically optimal molecular fingerprint structure under these conditions. Notably, attention mechanisms were introduced into both networks to enable the model to dynamically focus on different positions of the input molecular fingerprint sequence when generating output, rather than assigning equal importance to all input positions as with traditional fixed-weight models.[54] By learning attention weights, the model can decide which parts of the information to focus on based on the current partially generated output and the content of the input sequence, making the generated fingerprint structure more realistic and the discriminator's discrimination results more accurate. To address the potential issues of vanishing or exploding gradients as the number of network layers increases, residual modules were introduced into the network. The core idea of the residual module is to introduce a skip connection, or "residual connection," which directly passes the original input information to subsequent layers by adding the input signal to the output signal.[55] This skip connection facilitates the network's learning of the residual function, which is the difference between the original input and the target output, making it easier to optimize the model.

By calculating the similarity between the generated molecular fingerprints and each fingerprint in the database (eqn (5)), promising DESs were identified, where higher similarity indicated a higher possibility of high solubility for cobalt ions. To assess the accuracy and reliability of the CGAN model, DESs with high similarities were chosen for experimental validation. Additionally, ChCl : 6 *o*-cresol, a DES with low similarity, was randomly selected to ensure that DESs with low similarities truly exhibited poor leaching performance. As shown in Fig. 8(a), the overall trends of similarity and experimental results coincided with each other, with ChCl : Glutaric acid demonstrating the best performance and ChCl : 6 *o*-cresol exhibiting the lowest cobalt solubility. As descriptors for DESs'

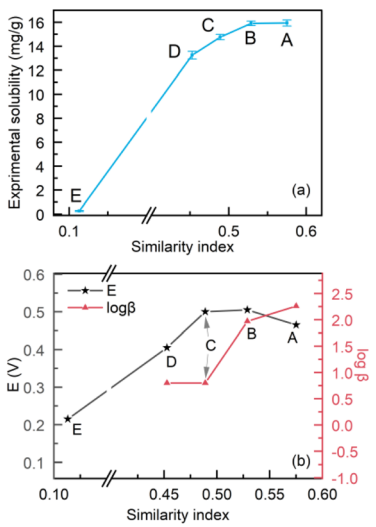

Fig. 8 (a) The correlation between the experimental outcomes and prediction results given (similarity index) by CGAN models. (b) The correlations of reaction potentials (E) and complex stability constants (log β) of DESs with similarity indices. The similarity indices are the results of the CGAN model on the DES database. Experimental solubilities of Co are the results of ICP-OES after leaching experiments in the conditions of T = 90 °C, time = 8 h, L : S = 30 g g^{-1}. The components and ratios of DESs (A–E) are as follows: ChCl : glutaric acid (A), ChCl : glycolic acid (B), ChCl : 3levulinic acid (C): trimethylamine hydrochloride : 3levulinic acid (D), ChCl : 6 *o*-cresol (E).

This journal is © The Royal Society of Chemistry 2024

Table 4 Comparison of leaching performance of different DES

DES	Temperature (°C)	L : S (g g^{-1})	Time (h)	Solubility of Co (mg g^{-1})	Ref.
ChCl : glycolic acid	90	30	8	15.93	This work
ChCl : glutaric acid	90	30	8	15.91	This work
ChCl : levulinic acid	90	30	8	14.76	This work
EG : sulfosalicylic acid	90	30	8	10.92	35
Betaine HCl : citric acid	100	50	1	4.0	56
Guanidine HCl : lactic acid	80	50	24	11.98	46
ChCl : L-ascorbic acid : H_2O	50	25	1	8.58	38
betaine HCl : EG	140	40	0.17	3.86	57

The experimental values in this work are the average of three replicates.

properties, E and $\log\beta$ were also used as conditions for the CGAN model, we utilized them to analyze the predicted results and get guidance for seeking promising DESs. Fig. 8(b) shows that the predicted results have similar trends with E and $\log\beta$, which reveals that the model could perform well by considering only two intrinsic properties of DESs. These results again underlined the importance of acidity coordination and reduction of DESs. Furthermore, to demonstrate the comparability of the predictions with the experimental works, we summarized the leaching conditions and performance of cathodes in different DESs from current studies to make a comparison. As shown in Table 4, the top three DESs have better performance than the others under comparable experiment conditions. In summary, these results confirm the reliability and potential of the CGAN model, demonstrating a promising application of ML in the discovery of favorable DESs for cathode recovery.

Conclusions

The machine learning models established in this work demonstrated their capability for accurate prediction, knowledge extraction, and expediting the discovery of promising DESs for spent cathode recovery. After the selection of an optimal feature set, the XGBoost model for cathode leaching in DESs exhibited excellent performance (R^2 = 0.9702, MSE = 0.0007), and the importance of coordination, acidity, and reducibility was quantified by SHAP interpretation. In addition, we took the MO dataset as an example to verify the generality of the proposed data analysis framework in the field of leaching metal oxides *via* DESs. Furthermore, a CGAN model was developed to rapidly identify promising DESs like ChCl : glycolic acid for leaching spent cathodes, showing excellent agreement with experimental results and saving significant time and effort compared to traditional trial-and-error methods. Therefore, ML could not only save the cost of research, but also promote sustainability by reducing chemical waste, unlocking new opportunities for more efficient and eco-friendly chemistry. Overall, the results highlight ML's promising application in discovering favorable DESs, deepening the understanding of leaching mechanisms and providing guidance to accelerate the development of spent LIB recovery. The proposed approach could advance the design and development of novel solvents for various applications.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank the National Natural Science Foundation of China (22238011, 22073112 and 72371014) for financial support.

References

1 W. E. Gent, G. M. Busse and K. Z. House, *Nat. Energy*, 2022, **7**, 1132–1143.
2 N. Muralidharan, E. C. Self, M. Dixit, Z. Du, R. Essehli, R. Amin, J. Nanda and I. Belharouak, *Adv. Energy Mater.*, 2022, **12**, 2103050.
3 C. B. L. Nkulu, L. Casas, V. Haufroid, T. De Putter, N. D. Saenen, T. Kayembe-Kitenge, P. M. Obadia, D. K. W. Mukoma, J.-M. L. Ilunga, T. S. Nawrot, O. L. Numbi, E. Smolders and B. Nemery, *Nat. Sustain.*, 2018, **1**, 495–504.
4 F. Maisel, C. Neef, F. Marscheider-Weidemann and N. F. Nissen, *Resour., Conserv. Recycl.*, 2023, **192**, 106920.
5 W. Mrozik, M. A. Rajaeifar, O. Heidrich and P. Christensen, *Energy Environ. Sci.*, 2021, **14**, 6099–6121.
6 G. Harper, R. Sommerville, E. Kendrick, L. Driscoll, P. Slater, R. Stolkin, A. Walton, P. Christensen, O. Heidrich, S. Lambert, A. Abbott, K. Ryder, L. Gaines and P. Anderson, *Nature*, 2019, **575**, 75–86.
7 P. Greim, A. A. Solomon and C. Breyer, *Nat. Commun.*, 2020, **11**, 4570.
8 H. Ji, J. Wang, J. Ma, H.-M. Cheng and G. Zhou, *Chem. Soc. Rev.*, 2023, **52**, 8194–8244.
9 J. Wang, K. Jia, J. Ma, Z. Liang, Z. Zhuang, Y. Zhao, B. Li, G. Zhou and H.-M. Cheng, *Nat. Sustain.*, 2023, **6**, 797–805.
10 J. Ma, J. Wang, K. Jia, Z. Liang, G. Ji, H. Ji, Y. Zhu, W. Chen, H.-M. Cheng and G. Zhou, *Nat. Commun.*, 2024, **15**, 1046.

11 J. Wang, Q. Zhang, J. Sheng, Z. Liang, J. Ma, Y. Chen, G. Zhou and H.-M. Cheng, *Natl. Sci. Rev.*, 2022, **9**, nwac097.
12 J. Wang, Y. Lyu, R. Zeng, S. Zhang, K. Davey, J. Mao and Z. Guo, *Energy Environ. Sci.*, 2024, **17**, 867–884.
13 Z. Yuan, H. Liu, W. F. Yong, Q. She and J. Esteban, *Green Chem.*, 2022, **24**, 1895–1929.
14 A. P. Abbott, D. Boothby, G. Capper, D. L. Davies and R. K. Rasheed, *J. Am. Chem. Soc.*, 2004, **126**, 9142–9147.
15 D. Yu, Z. Xue and T. Mu, *Chem. Soc. Rev.*, 2021, **50**, 8596–8638.
16 L. L. Chen, Y. H. Chao, X. W. Li, G. L. Zhou, Q. Q. Lu, M. Q. Hua, H. P. Li, X. G. Ni, P. W. Wu and W. S. Zhu, *Green Chem.*, 2021, **23**, 2177–2184.
17 M. K. Tran, M. T. F. Rodrigues, K. Kato, G. Babu and P. M. Ajayan, *Nat. Energy*, 2019, **4**, 339–345.
18 S. Wang, Z. Zhang, Z. Lu and Z. Xu, *Green Chem.*, 2020, **22**, 4473–4482.
19 M. J. Roldan-Ruiz, M. L. Ferrer, M. C. Gutierrez and F. del Monte, *ACS Sustainable Chem. Eng.*, 2020, **8**, 5437–5445.
20 H. Mai, T. C. Le, D. Chen, D. A. Winkler and R. A. Caruso, *Chem. Rev.*, 2022, **122**, 13478–13515.
21 P. Bannigan, Z. Bao, R. J. Hickman, M. Aldeghi, F. Häse, A. Aspuru-Guzik and C. Allen, *Nat. Commun.*, 2023, **14**, 35.
22 S. Tao, H. Liu, C. Sun, H. Ji, G. Ji, Z. Han, R. Gao, J. Ma, R. Ma, Y. Chen, S. Fu, Y. Wang, Y. Sun, Y. Rong, X. Zhang, G. Zhou and H. Sun, *Nat. Commun.*, 2023, **14**, 8032.
23 S. Tao, R. Ma, Y. Chen, Z. Liang, H. Ji, Z. Han, G. Wei, X. Zhang and G. Zhou, *J. Power Sources*, 2024, **597**, 234156.
24 M. Mohan, K. D. Jetti, M. D. Smith, O. N. Demerdash, M. K. Kidder and J. C. Smith, *J. Chem. Theory Comput.*, 2024, **20**, 3911–3926.
25 M. Mohan, O. N. Demerdash, B. A. Simmons, S. Singh, M. K. Kidder and J. C. Smith, *ACS Omega*, 2024, **9**, 19548–19559.
26 A. Zhu, X. Bian, W. Han, D. Cao, Y. Wen, K. Zhu and S. Wang, *Resour., Conserv. Recycl.*, 2023, **188**, 106690.
27 L. Grinsztajn, E. Oyallon and G. Varoquaux, presented in part at the NeurIPS 2022 Datasets Benchmarks, 2022.
28 D. Rogers and M. Hahn, *J. Chem. Inf. Model.*, 2010, **50**, 742–754.
29 A. K. Jain, M. N. Murty and P. J. Flynn, *ACM Comput. Surv.*, 1999, **31**, 264–323.
30 I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, *Commun. ACM*, 2020, **63**, 139–144.
31 J. Heaton, *Genet. Program. Evolvable Mach.*, 2018, **19**, 305–307.
32 P. Willett, J. M. Barnard and G. M. Downs, *J. Chem. Inf. Comput. Sci.*, 1998, **38**, 983–996.
33 H. Yu, S. Wang, Y. Li, Q. Qiao, K. Wang and X. Li, *Green Process. Synth.*, 2022, **11**, 868–874.
34 Q. Lu, L. Chen, X. Li, Y. Chao, J. Sun, H. Ji and W. Zhu, *ACS Sustainable Chem. Eng.*, 2021, **9**, 13851–13861.
35 S. Tang, M. Zhang and M. Guo, *ACS Sustainable Chem. Eng.*, 2022, **10**, 975–985.
36 K. Wang, T. Hu, P. Shi, Y. Min, J. Wu and Q. Xu, *ACS Sustainable Chem. Eng.*, 2022, **10**, 1149–1159.
37 G. Zeng, J. Yao, C. Liu, X. Luo, H. Ji, X. Mi and C. Deng, *ACS Sustainable Chem. Eng.*, 2021, **9**, 16133–16142.
38 Y. Hua, Y. Sun, F. Yan, S. Wang, Z. Xu, B. Zhao and Z. Zhang, *Chem. Eng. J.*, 2022, **436**, 133200.
39 B. Lu, R. Du, G. Wang, Y. Wang, S. Dong, D. Zhou, S. Wang and C. Li, *Environ. Res.*, 2022, **212**, 113286.
40 T. Li, Y. Xiong, X. Yan, T. Hu, S. Jing, Z. Wang and X. Ge, *J. Energy Chem.*, 2022, **72**, 532–538.
41 M. Liu, W. Ma, X. Zhang, Z. Liang and Q. Zhao, *Mater. Chem. Phys.*, 2022, **289**, 126466.
42 C. Ma, M. Svärd and K. Forsberg, *Resour., Conserv. Recycl.*, 2022, **186**, 106579.
43 F. Huang, T. Li, X. Yan, Y. Xiong, X. Zhang, S. Lu, N. An, W. Huang, Q. Guo and X. Ge, *ACS Omega*, 2022, **7**, 11452–11459.
44 Y. Chen, Y. L. Wang, Y. Bai, M. H. Feng, F. Y. Zhou, Y. H. Lu, Y. T. Guo, Y. X. Zhang and T. C. Mu, *Green Chem. Eng.*, 2023, **4**, 303–311.
45 Y. Chen, Y. Wang, Y. Bai, Y. Duan, B. Zhang, C. Liu, X. Sun, M. Feng and T. Mu, *ACS Sustainable Chem. Eng.*, 2021, **9**, 12940–12948.
46 Y. Tian, W. Chen, B. Zhang, Y. Chen, R. Shi, S. Liu, Z. Zhang and T. Mu, *ChemSusChem*, 2022, **15**, e202200524.
47 F. Zhou, R. Shi, Y. Wang, Z. Xue, B. Zhang and T. Mu, *Phys. Chem. Chem. Phys.*, 2022, **24**, 16973–16978.
48 L. Grinsztajn, E. Oyallon and G. Varoquaux, presented in part at the Proceedings of the 36th International Conference on Neural Information Processing Systems, New Orleans, LA, USA, 2024.
49 A. P. Abbott, G. Capper and S. Gray, *ChemPhysChem*, 2006, **7**, 803–806.
50 D. Shi, F. Zhou, W. Mu, C. Ling, T. Mu, G. Yu and R. Li, *Phys. Chem. Chem. Phys.*, 2022, **24**, 26029–26036.
51 A. Khawam and D. R. Flanagan, *J. Phys. Chem. B*, 2006, **110**, 17315–17328.
52 S. Tao, C. Sun, S. Fu, Y. Wang, R. Ma, Z. Han, Y. Sun, Y. Li, G. Wei, X. Zhang, G. Zhou and H. Sun, *ACS Energy Lett.*, 2023, **8**, 3269–3279.
53 S. H. Alhashim, S. Bhattacharyya, R. Tromer, A. Kabbani, G. Babu, E. F. Oliveira, D. S. Galvao and P. M. Ajayan, *ACS Sustainable Chem. Eng.*, 2023, **11**, 6914–6922.
54 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, *Advances in neural information processing systems*, 2017, p. 30.
55 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, presented in part at the Advances in neural information processing systems 30 (NIPS 2017), 2017.
56 Y. Luo, L. Ou and C. Yin, *Waste Manage.*, 2023, **164**, 1–8.
57 Y. Luo, C. Yin, L. Ou and C. Zhang, *Green Chem.*, 2022, **24**, 6562–6570.

This journal is © The Royal Society of Chemistry 2024