ADVANCED MATERIALS
www.advmat.de

Machine Learning: An Advanced Platform for Materials Development and State Prediction in Lithium-Ion Batteries

Chade Lv, Xin Zhou, Lixiang Zhong, Chunshuang Yan, Madhavi Srinivasan, Zhi Wei Seh, Chuntai Liu, Hongge Pan, Shuzhou Li, Yonggang Wen,* and Qingyu Yan**

Lithium-ion batteries (LIBs) are vital energy-storage devices in modern society. However, the performance and cost are still not satisfactory in terms of energy density, power density, cycle life, safety, etc. To further improve the performance of batteries, traditional "trial-and-error" processes require a vast number of tedious experiments. Computational chemistry and artificial intelligence (AI) can significantly accelerate the research and development of novel battery systems. Herein, a heterogeneous category of AI technology for predicting and discovering battery materials and estimating the state of the battery system is reviewed. Successful examples, the challenges of deploying AI in real-world scenarios, and an integrated framework are analyzed and outlined. The state-of-the-art research about the applications of ML in the property prediction and battery discovery, including electrolyte and electrode materials, are further summarized. Meanwhile, the prediction of battery states is also provided. Finally, various existing challenges and the framework to tackle the challenges on the further development of machine learning for rechargeable LIBs are proposed.

1. Introduction

With the increasing concerns on the environment and sustainability, it desperately demands advanced energy-storage technology to support new energy electric vehicles (EVs), and smart grids.[1] Attributed to the superior advantages of high operating potential and energy/power density, rechargeable lithium-ion batteries (LIBs) have emerged as a transformative technology since their development in the 1970s.[2] Nevertheless, the performance and cost are not satisfactory while the energy density of conventional LIBs have almost reached the theoretical maximum. To further advance current LIBs, numerous efforts have been devoted to the exploration of new electrode and electrolyte materials.[3]

The historical material research has heavily relied on either "trial-and-error" processes or serendipity, both of which require the vast numbers of tedious experiments (**Figure 1**a). Such intuition-based approaches are often time-consuming and inefficient, which cannot avoid the consumption of many manpower and material resources. In the past 50 years, computational chemistry, such as first-principles (FP) calculations,[4,5] quantum mechanics,[6] molecular dynamics (MD)[7] and Monte Carlo techniques,[8] has become a mature approach to complement and aid experimental studies for predicting and designing new materials. With the rapid development of high-performance computations, density functional theory (DFT) has been widely applied to high-throughput property prediction, which is conducive to the development of materials databases, such as Inorganic Crystal Structure Database (ICSD),[9] Cambridge Structural Database,[10] the Materials Project[11] database, AFLOWLIB consortium,[12] Open Quantum Materials Database,[13] Harvard Clean Energy Project,[14] Electronic Structure Project,[15] MaterialGo,[16] and so on. However,

C. Lv, L. Zhong, C. Yan, M. Srinivasan, S. Li, Q. Yan
School of Materials Science and Engineering
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798, Singapore
E-mail: lisz@ntu.edu.sg; alexyan@ntu.edu.sg

X. Zhou, Y. Wen
School of Computer Science and Engineering
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798, Singapore
E-mail: ygwen@ntu.edu.sg

M. Srinivasan, S. Li, Q. Yan
Energy Research Institute@NTU
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798, Singapore

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adma.202101474.

DOI: 10.1002/adma.202101474

Z. W. Seh
Institute of Materials Research and Engineering
Agency for Science
Technology and Research (A*STAR)
2 Fusionopolis Way, Innovis, Singapore 138634, Singapore

C. Liu
Key Laboratory of Materials Processing and Mold
Ministry of Education
Zhengzhou University
Zhengzhou 450002, China

H. Pan
Institute of Science and Technology for New Energy
Xi'an Technological University
Xi'an 710021, P. R. China

H. Pan
School of Materials Science and Engineering
State Key Laboratory of Silicon Materials
Zhejiang University
Hangzhou 310027, P. R. China

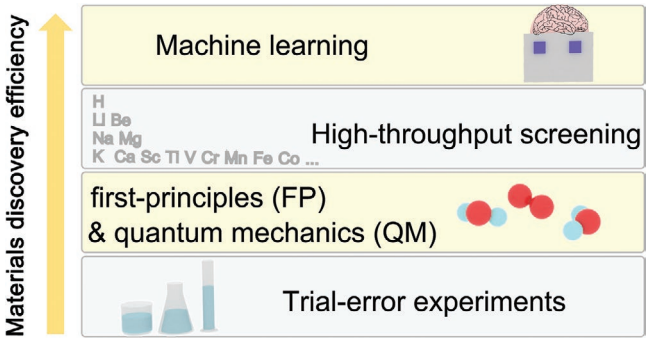


Figure 1. Development of methods in new materials discovery.

there are still many challenges in finding suitable descriptors or a model, which limit their effectiveness in exploring complex real-world materials due to the high computational cost and poor scaling of calculations.[17] For example, high-throughput DFT screening only allows limited search space (hundreds to thousands of materials), but mostly, the atomic number of materials is limited to less than 1000. Besides, large amounts of material information data are generally neglected in the databases.[18] Therefore, it is imperative to accelerate material innovation by finding a new research method.

Artificial intelligence (AI) is an emerging technique in material science.[19–22] Besides, the combination of AI with big data has been regarded as "the fourth paradigm of science".[23] Machine learning (ML), as the core part of AI, can find the statistical law behind high-dimensional data to produce reliable, repeatable decisions and results.[24] ML is also a viable technique to couple the small and large time-and-length scales.[25] With the ability of fast prediction of new materials or properties, it also guarantees high accuracy. ML models, including artificial neural network (ANN), support-vector machines (SVM), random forest (RF), partial least squares regression (PLS), and logistic regression (LR), have successfully predicted the properties of battery materials. For instance, Nakayama et al.[26] combined an ANN with DFT for the simultaneous prediction of diffusion barrier and cohesive energy (CE) of the candidate solid electrolyte material (olivine group $LiMXO_4$). Moreover, fifteen promising solid electrolyte materials for LIBs were screened out by using the same approach. Reed et al.[27] developed a data-driven ionic conductivity classification model using LR to identify the possible structures with fast lithium conduction. Twenty-one solid electrolyte materials were screened out from the MP database. Viswanathan et al.[28] conducted a computational screening of over 12 000 inorganic solids on the basis of their ability to inhibit the dendrite initiation on a Li-metal anode. Twenty mechanically anisotropic interfaces with four solid electrolytes were predicted to be able to suppress the growth of dendrite. Vegge et al.[29] presented a blueprint for reversely designing solid electrolyte interphase with excellent performance through utilization of semi-supervised generative deep learning models, high-throughput synthesis, and laboratory testing. Up till now, the publication number of academic articles has increased exponentially with the key words "AI, ML, and materials".[30] Undoubtedly, ML has become an effective computational method to acquire the composition–structure–property relationships in electrode materials with high efficiency and accuracy.[31–33]

As is known to all, the health and safety of LIBs is another concern.[34,35] In the last few years, the frequent fire accidents of EVs have driven the large demands for the battery-management system (BMS).[36–38] Therefore, developing advanced and intelligent BMS to accurately predicate the state of charge (SOC) and state of health (SOH) of batteries has become an important research topic. In essence, SOC is defined as the ratio of capacity in the current state to that in the fully charged state, while SOH reflects the current capability of a battery for energy storage relative to that when brand new.[39] Various models such as equivalent circuit models (ECMs), physics-based models (PBMs) have been proposed to on-line estimate the behaviors of batteries, hoping to obtain a precise SOC estimation.[40] Despite that, there still exists a distinct trade-off between the efficiency of calculation and the accuracy of model-based predictions.[41] Fortunately, ML models are able to predicate the state of battery since they have the excellent computational capability to handle any complex nonlinear function.[39,42,43]

Herein, we provide a heterogeneous category of AI technology for predicating and discovering battery materials and estimating the state of battery system. We also analyzed and outlined the successful examples, challenges of deploying AI in real-word scenarios, and an integrated framework. The remainder of the paper is organized as follows: We will first briefly provide three categories of ML or AI; then, we will summarize the state-of-the-art research about the applications of ML in property prediction and battery discovery, including electrolyte and electrode materials. Meanwhile, the prediction of battery states, such as SOC, SOH, and remaining useful life (RUL), is also provided in the above section. The last section discusses various existing challenges and the framework to tackle the challenges on the further development of ML for rechargeable LIBs.

2. Applications in Battery Territory with Heterogeneous Categories of AI Technologies

In this section, we present an alternative classification of ML algorithms, with respect to the AI capability and then briefly introduce the ML applications in LIBs. Subsequently, we outline the grand challenges of deploying ML technologies and propose a unified architectural framework to tackles these challenges.

Compared with the canonical classification (supervised, unsupervised, and reinforcement learning) reflecting the nature of the "teaching signals" that guide learning,[44] ML algorithms can be better categorized into the following three categories according to the AI capability for an application domain of interest:[45]

1) Descriptive AI relies on a set of historical data to yield insightful information and possibly prepare the data for further analysis. It is always used in data collection and analysis to thoroughly and deeply understand what happened in the system and can be potentially used to model system dynamics (e.g., simulator or digital-twin modeling for batteries);

2) Predictive AI attempts to understand the causes of system behaviors. It is able to predict the property and/or state of the system and is widely employed in fault diagnostics and anomaly detection to provide insight on the performance (e.g., battery property prediction);
3) Prescriptive AI provides ways to improve system efficiency. It derives optimized configurations/settings/actions for potential system design and management (e.g., battery design, material search, charging protocols, etc.)

This classification maps well with the well-accepted definition of AI techniques, for machines to approximate human intelligence, that is, understanding, reasoning, and ethics.[46] It follows that this classification is mutually exclusive collectively exhaustive, in that any algorithm has to fall under one and only one category.

The past decade has seen the rapid development of AI technologies, which accelerates the pace of LIBs research in material discovery and property prediction.[30,47] ML techniques can be used to construct correlations that exist in the data, where a crucial correlation in material science is the structure–property relation. ML-based applications in material science often begin with new dataset construction and/or existing dataset exploitation,[9,48] followed by some descriptive AI/ML approaches to extract correlations between energy efficiencies of different materials and structures that are then used as the structure–property selection rules.[49] Furthermore, the integrated approach of discriminative models accelerates the exploration of large chemical spaces with over 1.6 million molecules.[50] A predictive AI-based approach has been used to screen through 12 000 candidates, which renders the discovery of new Li-ion-conducting materials.[27] Unlike the discriminative models, deep generative models such as prescriptive AI approaches based on generative adversarial network[51,52] and reinforcement learning[53] can learn the mapping of conditional probabilities to model a new system itself, which enable materials inverse design. In the meanwhile, by thoroughly learning from training datasets, ML can model the conditional probabilities that predict certain properties/states given the necessary inputs. For instance, predictive AI/ML approaches based on regularized linear models and long short term memory models have achieved significant accuracy improvement for battery lifespan and temperature prediction.[41,54–56]

3. Applications of ML in Rechargeable LIB

Rechargeable LIBs have emerged as a revolutionary energy-storage technology, which underpins modern life. **Figure 2**a displays the reversibly shuttling lithium ions in electrolyte between two electrodes of the device. For improving the electrochemical performance, it is vital to develop suitable electrode and electrolyte materials (Figure 2b,c). Theoretically, we can discover the novel materials with the aid of predicting their properties. But generally, it is very difficult and expensive to determine the properties via standard DFT computational or experimental efforts. ML algorithms that can "learn" complex correlations and patterns from the existing data, provide a solution to the materials exploration problem. Meanwhile, although the screened battery materials may lead to better performance and more complicated battery dynamics, where the safety of batteries could be another concern. Thus, predicting the degradation behavior of batteries by ML techniques is also essential for the entire electrification system.[57–59] The fundamental goal of ML models in rechargeable batteries is to establish the quantitative structure–activity relationship (QSAR) between conditional attributes and decision attributes through low-cost and accurate predictions.[60] In this section, we will mainly focus on the recent applications of ML models for predicting properties of materials, state of battery, and designing materials for rechargeable LIBs.

3.1. Property Prediction

In the application of ML, the most prevalent form is property prediction, which is conducive for rapid materials screening. The basic workflow for battery material property prediction via ML methods as follows: first, feature engineering is beneficial to identify the conditional attributes; second, the mapping relationship is established between these conditional factors and the decision attributes through model training; last, various properties (battery voltage and ionic conductivity, etc.) can be predicted by the trained model. In the next subsections, we will discuss the details on properties prediction of rechargeable battery materials including electrode, liquid electrolyte, and solid electrolyte materials.

3.1.1. Electrode Materials

Seeking suitable electrode materials plays a significant role in the development of rechargeable LIBs. The properties of electrode materials, for instance, voltage, capacity, electronic/Li-ion conductivity, and chemical/electrochemical stability, have been taken into consideration for property prediction. Given that the intrinsic physical and chemical properties of electrode materials are determined by their crystal structure, it is essential to preferentially predict such properties. In this regard, for predicting three major crystal systems (monoclinic, orthorhombic, and triclinic) of silicate cathodes with Li–Si–(Mn, Fe, Co)–O compositions, Shandiz et al.[61] conducted a series of machine learning classification methods, including ANN, SVM, K-nearest neighbors (KNN), RF, and extremely randomized trees (ERT). Space group, formation energy, energy above the hull, band gap, number of sites, density, and volume of the unit cell served as descriptors, where RF and ERT methods realized the highest prediction accuracy. The accuracy was evaluated on the basis of Monte Carlo cross validation or called repeated random subsampling method. The result suggests the volume of crystal and the number of sites play the determinant role in the type of crystal system in the dataset. Moreover, as-proposed approach could provide better insight for other researchers to consider the correlations between various features of materials.

For the cathode materials, the intercalation characteristics of Li ions are pivotal for electrochemical performance. To achieve excellent performance, researchers are trying to understand the structure–performance relationships through regulating

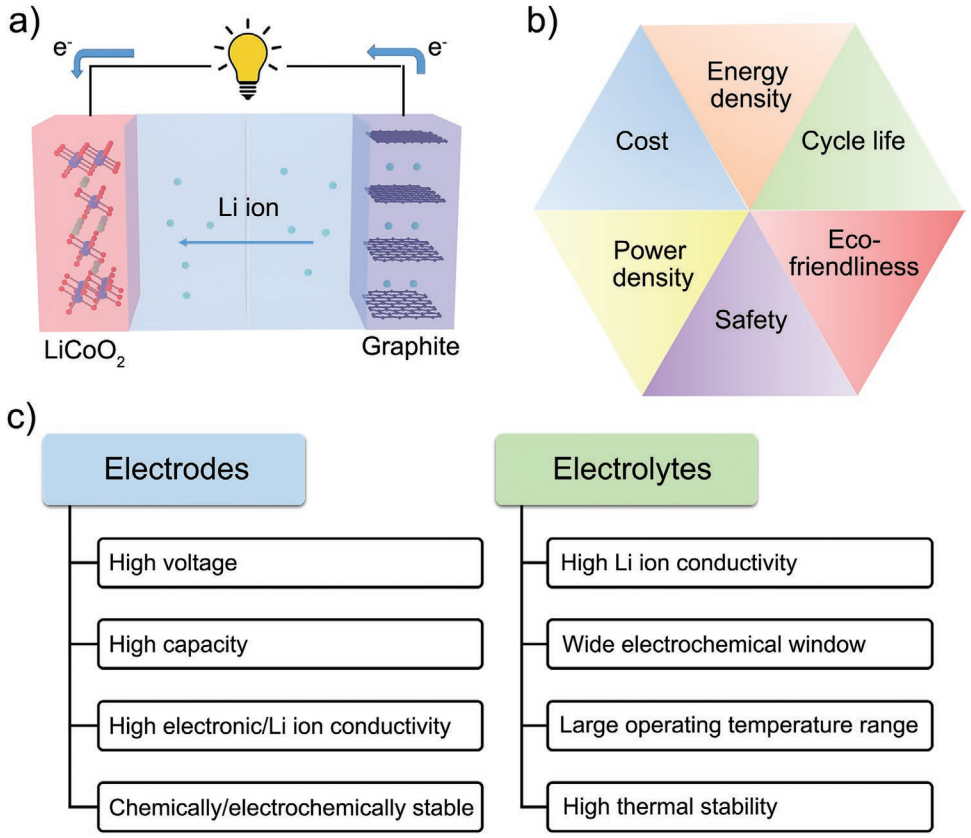


Figure 2. a) Schematic of the most commonly used Li-ion batteries based on $LiCoO_2$ cathodes and graphite anodes. b) The main properties of Li-ion batteries regarding performance. c) The desired properties of electrode and electrolyte materials.

the structures and elements of materials. As far as this is concerned, Xiao et al.[62] combined ab initio calculations and PLS analysis to develop the QSAR formulations of volume changes in spinel structure LiX_2O_4 and layered-structure $LiXO_2$, aiming at extending the cycle life of the cathode for LIBs. The authors noticed that the descriptors, the radius of X^{4+} ion and X octahedron, make a great contribution to the volume changes of the cathode. Impressively, the established QSAR formulation may be further employed to predict the volume change of the electrode in various real and "virtual" materials. The challenge lies in finding the optimal combination of parameters to lower the volume change for designing the low-strain cathode materials. Apart from the volume change for structural stability, the design of cathode materials with high voltage is also extremely desirable for achieving high energy density in LIBs. To this end, Sarkar et al.[63] utilized multilayer perceptron (MLP) ANN-based modeling to predict the voltage of different cathode materials for LIBs through choosing the central atom electronegativity and the stronger electronegative elements as input parameters. However, it lacks large data set to improve the accuracy of ANN model, which results in the main challenge in voltage prediction. On the basis of ML, Joshi et al.[64] exploited a tool to predict electrode voltages for metal-ion batteries. The deep neural networks (DNN), support vector regression (SVR), and kernel ridge regression (KRR) were applied as ML algorithms in combination with taking data from the Materials Project database, together with feature vectors from properties of chemical compounds and elemental properties of their constituents. Using their models, nearly 5000 electrode materials were proposed as candidates for Na- and K-ion batteries. It might be essential to further boost the performance of the model for the routine application of ML algorithms in predicting the voltage of electrode materials. Eremin et al.[65] applied ridge regression (RR) method to predict the energy of $LiNiO_2$ and $LiNi_{0.8}Co_{0.15}Al_{0.05}O_2$ cathode materials. They found that the topology of Li layers and relative disposition of Li and Al in the structure of $LiNi_{0.8}Co_{0.15}Al_{0.05}O_2$ are the most critical descriptors during the energy balance estimations. Moreover, to find the link between manufacturing parameters and macroscopic properties of electrodes, Franco et al.[66] proposed an AI-based computational strategy to predict the role of the manufacturing parameters (active material mass loading and porosity) in the characteristics of the final electrode. Three different ML algorithms, including decision tree (DT), DNN, and SVM were also tested by Franco et al.[66] Among them, SVM accurately uncovered several trends linking the electrode mass loading and porosity to the slurry characteristics. Based on the abovementioned reports, many advanced ML algorithms were conducted to predict the key parameters of cathode materials with high accuracy. In addition, the SVM method, a kernel-based regression technique known for its robust performance in complex data representations,[64] was frequently carried out for predicting the cathode materials.

As to anode materials, the ANN approaches were usually employed because it enabled implicit classification of complex

nonlinear relationships among a relatively large number of variables.[67] Allam et al.[68] tried to establish a DFT-ML framework to predict the redox potential of carbon-based molecular electrode materials. Based on this analysis, it suggested that the electron affinity possessed the highest contribution to redox potential, followed by the number of oxygen atoms, the HOMO–LUMO gap, the number of lithium atoms, LUMO, and HOMO in order, respectively. The predicted redox potentials through ANN were in accordance with those calculated from DFT modeling, whose averaged error was 3.54%. ANN, a valid theoretical tool, was developed by Kalaiselvi et al.[69] to understand the charge–discharge characteristics of CoO anode. The experimental capacity values well agreed with the estimated/predicted ones. The chosen ANN approach realized the best fit values with an error factor of <1%. Thanks to the as-employed ML method, high-accuracy predictions were realized for redox potentials or charge–discharge behavior, which were significant parameters for anode materials in LIBs.

To summarize, it indicates the successful prediction of electrode materials need to combine the proper descriptors with ML methods based on the above examples. Some embedded FS methods (such as RF or ERT), as well as some correlation analysis methods (such as contribution analysis or sequential backward selection algorithm), were carried out to acquire the key factors, which could affect the properties of electrode materials. Attributed to this, novel materials would be rationally designed by researchers. However, in the majority of cases, the relationship between the chosen descriptors and targeted properties of materials is complex and nonlinear. Thus, these ML models require large data sets to train on. To this end, the employment of multiple ML algorithms for modeling and the generation of big data obtained by the virtual simulation of the digital twin might be beneficial for the optimal prediction ML model.

3.1.2. Electrolyte Materials

Apart from electrode materials, electrolytes, typically including the liquid and solid ones, are also indispensable components of LIBs. Recently, ML-based methods have provoked numerous attentions of many researchers on the properties prediction of electrolytes. The current studies are mainly focused on the solid electrolyte. However, it still remains a dilemma to estimate the properties of liquid electrolytes due to the difficulty in extracting information from disordered structures. For Liquid electrolytes, the ion transport at the electrolyte/electrode interface is always concerned and has a large impact on the rate performance. Generally, the property of the Li^+–solvent pair plays a key role in studying the ion transport phenomenon. For instance, the coordination energy (E_{coord}) and melting points of solvents were predicted by Sodeyama et al.,[70] who also discussed the extracted descriptors through three techniques, including multiple linear regression (MLR), least absolute shrinkage and selection operator (LASSO), and exhaustive search with linear regression (ES-LiR). Meanwhile, the estimation accuracy of the three techniques was examined in seeking liquid electrolyte materials. Among above techniques, ES-LiR reproduced the highest accuracy in the properties estimation. They also found that ES-LiR could establish the relationship between the "'prediction accuracy"' and "'calculation cost"' of the properties through a weight diagram of descriptors. Similarly, Ishikawa and Sodeyama et al.[71] applied an ML-based technique, in combination with quantum chemistry calculations, to obtain an accurate and efficient approach for predicting the E_{coord} values of the ions to the solvent. The E_{coord} of alkali metal ions to solvents was first calculated by DFT for Li, Na, K, Rb, and Cs ions and 70 solvents. Then the calculated E_{coord} was employed as the target properties in the regression using MLR, LASSO, and ES-LiR methods. They found that the ionic radius was the most significant descriptor and ES-LiR could provide the high accuracy for the prediction of E_{coord}. ES-LiR applies exhaustive search to MLR and introduces the indicator that represents a combination of non-zero explanatory variables. It is possible to estimate the variable and seek the proper indicator that optimizes the combination of descriptors when the relationship between descriptors and the objective variable is linear. However, ES-LiR may fail when the relationship is not linear. To learn the nonlinear relationship and obtain higher accuracy, the algorithm that applies exhaustive search with Gaussian process is ES-GP. The efficiency and accuracy for practical use in seeking for battery electrolytes were high enough with as-proposed regression models.

In recent years, the use of electrolyte additives as an economical way is expected to promote the electrochemical performance. The redox potential of additive is one of the most important indicators that allow their use as additives. Okamoto et al.[72] conducted ab initio molecular orbital calculations to investigate the redox potentials of 149 representative molecules that could serve as electrolyte additives. Then the calculated potentials were trained to build regression models by employing ML-based method. They chose the chemical structures of additive molecules as descriptors to predict the redox potentials by the Gaussian kernel ridge regression (GKRR) and gradient boosting regression (GBR) methods. Although the two methods well reproduced the oxidation potentials, GBR showed the superiority in predicting the reduction potentials. GKRR combines the Gaussian kernel method and RR with L2-norm regularization term. The Gaussian kernel provides nonlinear character and can efficiently learn relationships that traditional linear regression fail to obtain. In the meanwhile, GBR applies ensemble DT to build a regression model so that it can be updated by minimizing the loss function according to the gradient in a stepwise manner. Therefore, when there are sufficient training data, GBR will outperform GKRR approach. It is worth noting that a principal cause of LIBs failure is the degradation of electrolyte. Thus, it is very important to ascertain the concentration of $LiPF_6$ and weight fractions of solvents in an unknown electrolyte. Normally, the electrolyte is analyzed quantitatively by employing gas chromatography–mass spectrometry (GC-MS), inductively coupled plasma optical emission spectrometry (ICP-OES), and et al, while these instruments are high-cost and the corresponding measurements are complex. Dahn et al.[73] presented a new method by using Fourier transform infrared and an ML algorithm to probe the concentrations of major components in the liquid electrolytes. The method agreed well with the consequences from GC-MS/ICP-OES, which was favorable for accelerating electrolyte analysis of aged LIBs and disclosing the cell degradation mechanisms.

In summary, different properties of liquid electrolytes, such as temperature, molecular concentration, composition, have served as descriptors for prediction. Due to the complex nonlinear relationship between the descriptors and the properties of electrolytes, SVM and ANN were usually employed. In abovementioned reports, the small data restricted the ANN model, of which the prediction performance would be enhanced when increasing the sample size.

All-solid-state batteries have drawn great interest in recent years. Using solid electrolytes is a promising solution for solving the safety issues, improving energy density and widening electrochemical window of LIBs.[74] The desired solid electrolyte materials need to satisfy several criteria, such as fast Li-ion conduction, wide electrochemical window, negligible electronic conduction, high mechanical rigidity, and robust chemical stability. Many ML-based methods have been reported to predict these properties of solid electrolytes. For instance, Nakayama et al.[26] developed a supervised feed-forward neural network (NN) to predict the Li-ion diffusion barrier/hopping energy (EA) and CE in the olivine-type $LiMXO_4$ solid electrolytes. In the model, several structural and component variables are identified as key descriptors such as the intra-atomic parameters, effective charges of cations, ionic radius, and so on. Impressively, the model with two external attributes is more accurate and interpretable in contrast to the model with a single response variable. The feed-forward NN model indeed shows the better prediction result for the Li-ion diffusion barrier as compared to PLS-derived models.[75] By combining DFT computation with NN, they construct a model to predict migration energy (ME), where input features are derived from the database fulfilled by structural information from DFT-optimized crystal structures of various compositions. The input features are preprocessed and fed into the NN model for the training and validation. Subsequently, the model randomly learns many local atomic environments and then is able to predict unknown structure types that are not included in the training dataset. Moreover Nakayama et al.[76] also succeeded in formulating the feed-forward NN model to predict the ME of tavorite-type $LiMTO_4F$ solid electrolytes with the same descriptors. Theoretically, the selected predictor variables in the feed-forward NN model can be applied to the other structure types. Later, Nakayama et al.[77] established a Bayesian-optimization-driven DFT-based approach to predict the ME of tavorite-type Li- and Na-containing compounds. In sharp contrast with random search even under a strict condition of having a positively skewed ME sample distribution, the Bayesian optimization search method was relatively more efficient.

Apart from the ME and EA, the Li-ion conductivity also can be directly predicted by using ML-based methods. For example, Ibrahim et al.[78] studied the impact from chemical composition and temperature on the ionic conductivity of the polymer electrolyte system. A Bayesian neural network (BNN) was developed to predict the ionic conductivity of the nanocomposite polymer electrolyte system. In the BNN model, the inputs are chemical compositions and temperatures, while the outputs are the ionic conductivities of the polymer electrolytes. The predicted results demonstrated that different chemical compositions and temperatures affected the ionic conductivity of the polymer electrolyte system, in line with the experimental results. Tanaka et al.[79] employed SVR method with a Gaussian kernel to predict the low-temperature ionic conductivities of the LISICON-type superionic conductors with descriptors of diffusivity at 1600 K, the average volume of the disordered structure, order–disorder phase transition temperature. These descriptors were determined from the theoretical and experimental data. Using the model, the authors predicted that γ-Li_4GeO_4 and several other compounds exhibited a few time higher ionic conductivity than LISICON $Li_{3.5}Zn_{0.25}GeO_4$ at 373 K. Albeit only pseudobinary solid solutions being considered in the study, the as-presented methodology was not strictly confined to such systems, indicating its potential for rational prediction of other Li-ion conductors. For garnet-type solid electrolytes, Kireeva et al.[80] applied SVR model to predict the Li-ion conductivity, thus developing the composition–structure–ionic-conductivity relationships and surveying garnet-related structures for promising compositions with *t*-stochastic triplet embedding. The data visualization techniques are attractive for virtual screening. Miwa et al.[81] studied the Li-conduction properties of Nb-doped garnet-type oxide $Li_7La_3Zr_2O_{12}$ (LLZO) by the MD simulations with the MLP. The predicted Li-ion conductivity at 298 K and the activation energy matched well with the experimental data. Furthermore, the as-presented approach correctly predicted two Li occupation sites of 24 d and 96 h, offering the 3D network of the Li migration pathway.

The prediction of mechanical properties for solid electrolytes is also very imperative because the dendritic growth of Li metal anodes can be inhibited by them. The crystal graph convolutional neural network (CGCNN) model was proposed by Ahmad et al.[28] to predict the shear and bulk moduli of the crystalline solid electrolyte materials. Besides, they employed GBR and KRR to predict the elastic constants of materials with the cubic crystal structure. As the core part of the CGCNN model, the multigraph representation of the crystal structure encoded the atomic and the bonding interactions information. As a result, the authors identified that the promising solid electrolytes with some common features such as high anisotropy and mechanical softness could optimize the dendrite suppression and ion conduction. Liu et al.[82] applied SVM and KRR models to estimate the possible reactions and thermodynamic stability of Li|$Li_7La_3Zr_2O_{12}$ (LLZOM, M = dopant)) interfaces. Through the as-proposed model, they predicted 18 unexplored dopants M in LLZOM systems against Li metal, which were affirmed in the automated route built by an FP approach. According to their study, the M—O chemical bond strength plays a dominant role in the stability of the Li|LLZOM interface. Assisted by machine learning, Hatzell et al.[83] employed advanced in situ imaging techniques to track morphological transformations at Li|LLZO interface. The ML methods realized segmentation of lithium and pores from the reconstructions of in situ conditions. The ML-assisted in situ X-ray imaging technique offered data for physical insight into microstructure transformation in Li metal as well as solid electrolytes during cycling. Based on the above examples, when applying descriptors such as ion radius, electronegativity, bond lengths, and bond angles, it enabled ideal prediction in Li-ion conductivity for solid electrolyte materials.

Albeit differences existing in the various algorithms, standout prediction performance for electrode materials and electrolytes has been achieved by most of the ML methods, as summarized in **Table 1**. However, the data availability is the major limitation for the applications of ML in materials prediction.

Table 1. Application of ML in the prediction of rechargeable battery materials.

Material types	Materials	Dataset	Algorithms	Achievements	Refs.
Cathode	Li–Si–(Mn, Fe, Co)–O	Materials Project, ICSD	ANN, SVM, KNN, RF, ERT	Predicting 3 major crystal systems	[61]
Cathode	LiX_2O_4 (X = Ti, V, Cr, Mn, Fe, Co, Ni, Nb, Mo, Ru, Rh, Pd, Ta, Ir)	Calculated results from DFT	PLS	Predicting the volume change values of 28 cathodes	[62]
Cathode	$LiNiO_2$, $LiMn_2O_4$, $LiTiS_2$, $LiV_2(PO_4)_3$, etc.	ICSD	ANN	Predicting the voltage values of 31 cathodes	[63]
Cathode	$ACoO_2$, $ANiO_2$, $ATiO_2$, etc. (A = Li, Na, K)	Materials Project	DNN, SVR, KRR	Proposing nearly 5000 candidate electrode materials for Na- and K-ion batteries	[64]
Cathode	$LiNiO_2$, $LiNi_{0.8}Co_{0.15}Al_{0.05}O_2$	Computed statistics	RR	Predicting the energy of configuration in LNO and NCA cathodes	[65]
Cathode	$LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$	ARTISTIC project webpage	DT, DNN, SVM	Predicting the impact of the manufacturing parameters on the final electrode characteristics	[66]
Anode	Carbon-based molecular electrode materials	Data from other publications	ANN	Predicting the redox potentials with low averaged error	[68]
Anode	CoO	Experimental data	ANN	Predicting the charge–discharge behavior of CoO with an error factor of <1%	[69]
Liquid electrolyte	103 solvent molecules, such as propylene carbonate and ethylene carbonate	103 solvent molecules which were commercialized as battery-grade materials from KISHIDA Chemical Co., Ltd.	MLR, LASSO, ES-LiR	Predicting the coordination energies and melting points of 103 solvent molecules	[70]
Liquid electrolyte	70 solvents and five ions (Li, Na, K, Rb, and Cs)	70 solvents taken from commercialized battery-grade materials from KISHIDA Chemical Co., Ltd.	MLR, LASSO, ES-LiR	Predicting the coordination energies of 70 solvents	[71]
Electrolyte additives	149 electrolyte additives	Data from other publications	GKRR, GBR	Predicting the redox potentials of 149 representative molecules	[72]
Solid electrolyte	$LiMXO_4$ (M—main group elements, X—group XIV and group XV elements)	The chemical search space in the $LiMXO_4$ system, a total of 72 compositions	NN	Predicting the Li-ion diffusion barrier and CE of 72 compositions	[26]
Solid electrolyte	Olivine-type $LiMXO_4$ (main group M^{2+}–X^{5+}, M^{3+}–X^{4+})	Li-ion hopping energy computed by nudged-elastic-band method	PLS	Predicting the Li-ion hopping energy combined with nudged-elastic-band method	[75]
Solid electrolyte	Tavorite-type $LiMTO_4F$	Database of structural information extracted from DFT optimized crystal structures of different compositions (M–T pairings).	NN	Predicting the Li ME of $LiMTO_4F$ with chemical substitutions at M and T sites	[76]
Solid electrolyte	Tavorite-type $AMXO_4Z$ (A: Li, Na; M: group 2, 3, 4, 13 elements; X: group 14, 15, 16 elements; Z: F, Cl, Br, I)	Data from other publications and newly calculated datasets	Bayesian-optimization-driven DFT-based approach	Predicting the ion ME of tavorite-type Li- and Na-containing compounds	[77]
Solid electrolyte	Nanocomposite polymer electrolyte system (PEO–$LiPF_6$–ECCNT)	Experimental data	BNN	Predicting the ionic conductivity of nanocomposite polymer electrolyte systems (PEO–$LiPF_6$–EC–CNT)	[78]
Solid electrolyte	LISICON-type superionic conductors	Theoretical and experimental data	SVR with a Gaussian kernel	Predicting the low-temperature ionic conductivity of 72 compounds	[79]
Solid electrolyte	Garnet-type metal oxides	Experimental data	SVR	Predicting the Li-ion transport characteristics and identifying the descriptors that are responsible for high Li-ion conductivity in garnet-structured oxides	[80]
Solid electrolyte	Garnet-type oxide $Li_7La_3Zr_2O_{12}$ (LLZO)	Theoretical and experimental data	MD simulations with (MLP)	Predicting the Li-conduction properties of Nb-doped garnet-type oxide $Li_{6.75}La_3(Zr_{1.75}Nb_{0.25})O_{12}$	[81]

Table 1. Continued.

Material types	Materials	Dataset	Algorithms	Achievements	Refs.
Solid electrolyte	12 000 inorganic solids	Materials Project	CGCNN, GBR, KRR	Screening 12 950 solids using isotropic stability criteria and over 15 000 interfaces using anisotropic stability criteria of electrodeposition on the Li metal anode	[28]
Solid electrolyte	$Li_7La_3Zr_2O_{12}$ (LLZOM, M = dopant)	Materials Project	SVM, KRR	Evaluating possible reactions and the thermodynamic stability of Li|LLZOM interfaces under various chemical conditions	[82]
Solid electrolyte	LLZO	Experimental data	Deep convolution neural network	Effective segmentation to extract quantitative metrics of the electrodes during cycling	[83]

3.1.3. Battery State Prediction

LIBs are increasingly playing a pivotal role in accelerating the electrification process of vehicles. For EV batteries, there are five crucial guidelines: lifetime, specific energy, specific power, cost, and safety.[84,85] The former four aspects have achieved great improvements by optimizing the electrode and electrolyte materials during the past decade,[86] however, the safety issue has not been adequately addressed by most of the stakeholders in the EV market. Accurate determination of SOC/SOH and reliable prediction of RUL will alleviate the problem and unlock improvements in battery manufacturing, usage, and optimization. For example, end users can estimate the expected battery life to allow the usage of batteries to their fullest potential before replacement or disposal.[87–89] Likewise, manufacturers can grade new cells by their expected lifetime to accelerate the cell testing, validation, and manufacturing processes. Thus, an intelligent BMS that can predict and monitor battery behavior is essential for the entire electrification system.[39,90,91]

In order to develop an intelligent BMS, battery modeling, as the core part, is vital in determining the current state of battery and predicting the future state of batteries.[92–95] The battery models studied in the literature mainly fall into the empirical/semi-empirical models, ECMs, PBMs and, more recently, data-driven models (DDMs) with artificial intelligence algorithms (**Figure 3**).[96–98] Each model has its own advantages and drawbacks with respect to accuracy and complexity. For example, many empirical and semi-empirical approaches are very simple, but the predictive power is lost due to the oversimplification in some cases.[99] ECMs have gained much interest in real-time status predictions due to their simplified model structure. But it remains a grand challenge in obtaining high accuracy.[100] PBMs can provide the internal physical and chemical properties of batteries such as the Li-ion concentration, however, it is difficult to be applied in real-time applications due to the computational complexity of the coupled partial differential equations and a large number of unknown variables.[101] Based on the above, an appropriate balance between model fidelity and computational complexity has become the key obstacle in current battery models. Recently, DDMs with ML techniques have emerged as a potential modeling approach since they had the excellent computational capability to handle any complex nonlinear functions with low computational cost.[102–104] Generally, ML uses a fitting function from the experimental training data to make predictions for other battery systems. Various models such as linear model,[54,105] ANNs,[106–113] SVM,[114–120] RF,[121–123] Kalman filters,[124–126] gated recurrent unit recurrent neural network,[127] convolutional neural network,[128,129] DNN,[130] JAYA,[131] metabolic extreme learning machine,[132] and Gaussian/Bayesian regression,[133–135] have been reported to be able to predict the states of batteries. Below we will elaborate some of the most recent DMMs employed to estimate the different battery states for LIBs. For example, Severson et al.[54] used a linear regression model to predict the cycle life of

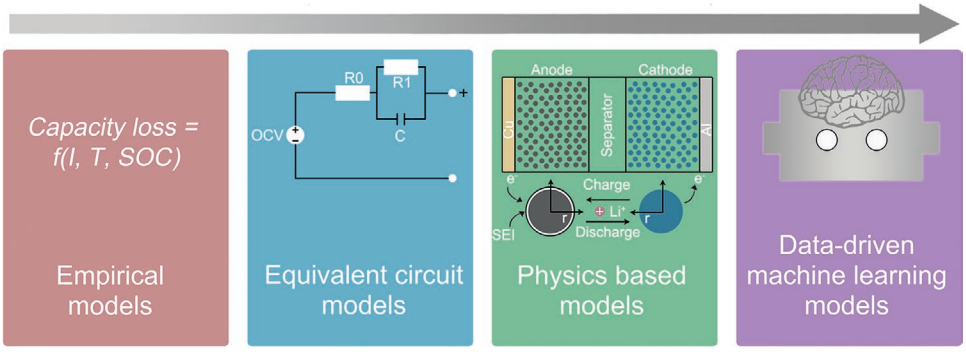


Figure 3. Development of battery models.

commercial lithium iron phosphate/graphite batteries through statistical and machine-learning techniques. The simple linear model with a typical error of 9.1% allows low computational cost and fast computational time. Moreover, Li et al.[105] also found a linear regression relationship between the specific peak locations from the incremental capacity analysis spectra and the SOH of a cell. The developed estimation function enabled less than 2.5% maximum errors in the SOH evaluation of batteries.

However, the linear models are unstable for RUL estimation since the battery capacity fade may accelerate under extreme conditions.[136] To solve this, Fermin et al.[137] proposed the identification mechanism of "knee-onset" (beginning of the nonlinear degradation) and "knee-points" by using a combination of a Bacon–Watts model and an SVM. The RF machine is successfully used to predict the RUL of the battery.[138] The approach enabled a typical prediction error of 3.3% by only reading the variation of voltage with the time. However, it is difficult to ensure the accuracy of the model when the RF algorithm is utilized for the prediction of battery life, leading to a big prediction error. In view of the issue, Li et al.[121] proposed to use the genetic algorithms to optimize the RF model, which strengthens the prediction accuracy. An SVM was employed by Nuhic et al.[139] to simultaneously predict SOH and RUL of LIBs. Taking account of the influence of environmental changes and loading conditions, the SVM was combined with a new method for training and data testing based on loading collectives. Yang et al.[140] proposed a three-layer back propagation NN to evaluate the SOH with a 5% error. The parameters of the first-order ECMs served as inputs, while the current value of SOH was the output. Zahid et al.[141] presented a subtractive clustering-based neuro-fuzzy system to estimate SOC of LIBs. The model inputs current, temperature, actual power dissipation, and available power to predict the SOC with a maximum estimation error < 0.1%. Experimental results illustrated that the proposed model enabled sufficient accuracy and surpassed the NN model. Sahinoglu et al.[134] employed Gaussian process regression framework to estimate the SOC of LIBs. The measured battery parameters, such as voltage, current, and temperature, were used as inputs. The simulations and experimental results suggested the advantage in comparison to the SVM and NN predictions.

Among these models, it is a complicated problem to select an appropriate ML approach, which relies on the amount of data available, the expected quality of results and the physical interpretability of the model. Due to the high levels of accuracy when predicting SOC, NNs are the most preferable approach. Nevertheless, the preferred ML approach is more nuanced during the SOH or RUL prediction. For example, Gaussian processes were employed[142–144] owing to the relative insufficiency in data, which could lead to probably safety-critical health diagnostics. The RUL could be represented as the number of remaining charge/discharge cycles. The integer rather than continuous quantity rendered it suitable for RF. Moreover, as many ML approaches were black boxes, it was more critical for a physical understanding. This manifested that it was better for the straightforward nature of linear regression.

Although great efforts have been made in developing data-driven ML approaches for battery management, there are still several challenges. For instance, most studies mainly rely on the collected data from a small number of cells, which is often not shared. Thus, it lacks systematically collected, standardized, and accessible experimental battery data. Recently, an abundance of data from the internet-of-things becomes available attributed to the low-cost sensing and increased deployment of devices.[145] Another challenge is to comply with the constrains of real-world deployment when constructing data-driven ML models by designing battery cycling experiments. Nevertheless, these experiments are high cost and time consuming, which is not feasible for small, individual laboratory. The digital-twin approach could establish a virtual representation of the physical system to simulate the variable real-world operating conditions. Besides, the approach can collect enough data to improve the ML models.[145] More details will be discussed in Section 4.

3.2. Materials Discovery and Design

The discovery and design of new battery materials aims to find proper electrodes or electrolytes with desirable properties to improve the performance and safety of LIBs. The modern computational tools are able to make a prediction for the properties of particular materials under specific conditions. However, up to now, the inverse design of battery materials has been computationally infeasible due to the massive complexity.[146,147] As mentioned in section 3.1, the construction of a model is the most vital step in the property prediction, which will precisely depict the relationship between the input descriptors (structural or elementary information) and output target properties (conductivity or stability) of the known materials. Contrary to the property prediction, the properties of materials are the input and the structure and composition are the output in an inverse materials design. That is, the key issue is to identify the chemical components and structures of materials, which can be synthesized in the lab.

For the discovery and design of materials, the first step is to generate the key descriptors or features that are closely associated with the material property of interest (**Figure 4**). The construction of an accurate model between the descriptors and target properties is the second step. Theoretically, on the basis of the ML model trained in a given dataset (materials→properties), the inverse design can be conducted to discover the new materials with the intended properties. Two major approaches are involved-large scale screening and mathematical optimization to realize the design process. For the large scale screening method, it is the first step for the generation of all possible target materials in the design space, followed by the test of materials using the as-built ML model.[148] Meanwhile, there are some constraints to consider on the material representation in the form of a structure or composition-based function. In terms of the issue, a systematic process is necessary to identify these candidate materials. On the other hand, the reverse materials design can also be formulated as a mathematical optimization problem. The optimization-based method tries to identify the candidate materials without testing them one by one, reducing the complexity.[149,150] Once the optimal materials are identified, one can synthesize them and verify their properties experimentally. If the experimental results are consistent with the predicted ones computationally, the materials are discovered

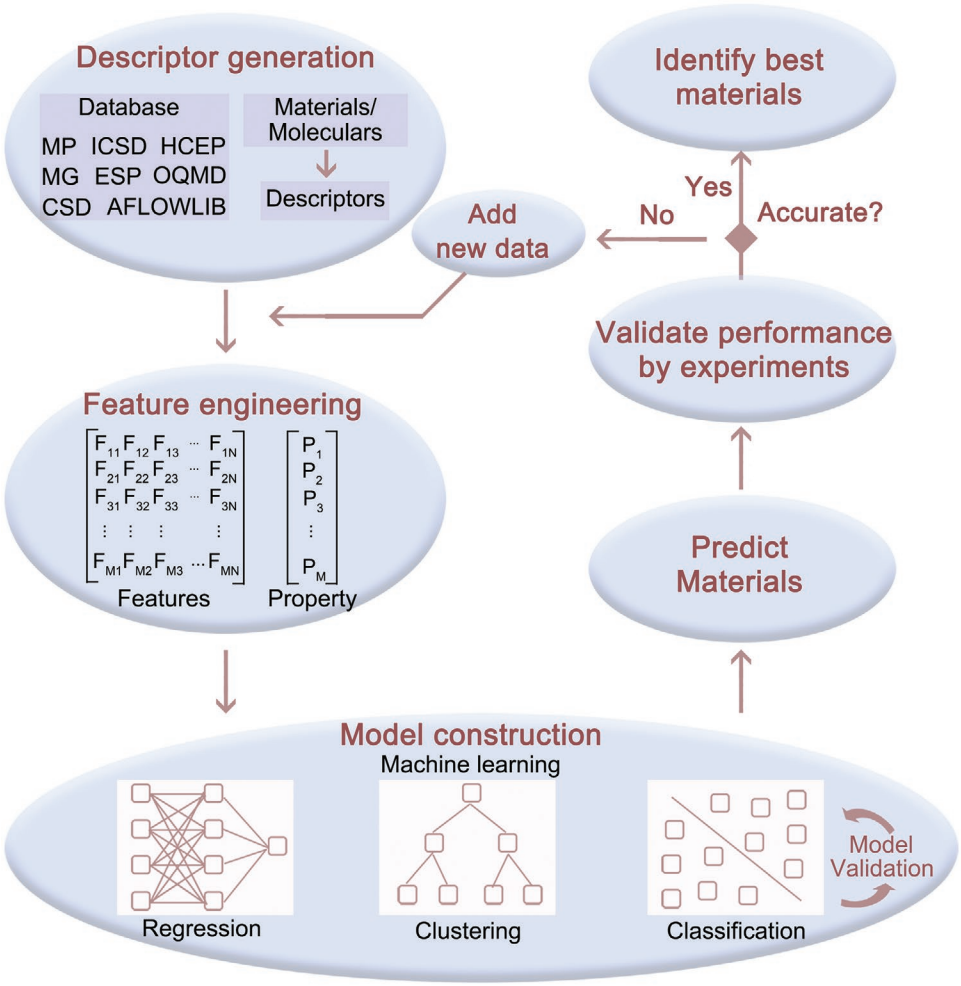


Figure 4. The basic workflow for battery materials discovery and design by ML methods.

and designed successfully. If not, the obtained results can be added to the training dataset to retrain the ML model. The cycle creates a feedback loop to improve the accuracy and refine the followed discovered and designed cycles. Although the computational materials discovery has made great progress that was outlined in the 2019 materials design roadmap,[151] it is short of successful examples of inverse materials design and the closed-loop inverse materials design will be the long-term goal.

Next, some selected examples are analyzed to highlight the role of ML in discovering battery materials. As we all know, the high-capacity Li-excess $LiNi_xMn_yCo_{(1-x-y)}O_2$ (NMC) layered oxides have been considered as a promising candidate because of their high reversible capacities. However, the practical application is limited due to the severe voltage decay. A key driver for improvements is the modifications of the cathode composition. Houchins et al.[152] developed a NN potential using fingerprinting and performed hyperparameter optimization of the fingerprinting parameters. They use the ML calculator to predict the structural effects due to insertion/de-insertion of lithium and the open-circuit voltage for any composition of $LiNi_xMn_yCo_{(1-x-y)}O_2$ cathode. The predicted voltage profiles are highly consistent with the experimental ones. This provides an approach to rapidly design and optimize the NMC cathode material family in the phase space. The performance of LIBs can be also improved by introducing interfacial coating materials, which can eliminate the formation of undesired interphases and enhance the cyclability of batteries. Wang et al.[153] employed the machine-learned interatomic potentials models in the form of moment tensor potentials to identify two promising coating materials. This approach reduces the computing time and increases the efficiency of the calculations by 7 orders of magnitude relative to the pure ab initio MD (**Figure 5**a). The redox stability of electrolyte also has an enormous impact on the electrochemical performance. For this, one can design new organic solvents with proper redox potential possessing reduction and oxidation stability at both the anode and cathode.[154] Tagade et al.[155] proposed a binary representation to digitize the molecular structure and used the semi-supervised algorithm to map the relationship between structures and properties. Meanwhile, they applied Bayesian approach to ascertain the generation of the valid solvent molecular structures. Finally, many organic solvent structures were discovered, which had a lower reduction potential than the used anode to enable an electrochemical window of 4.8 V. These desired attributes guarantee the

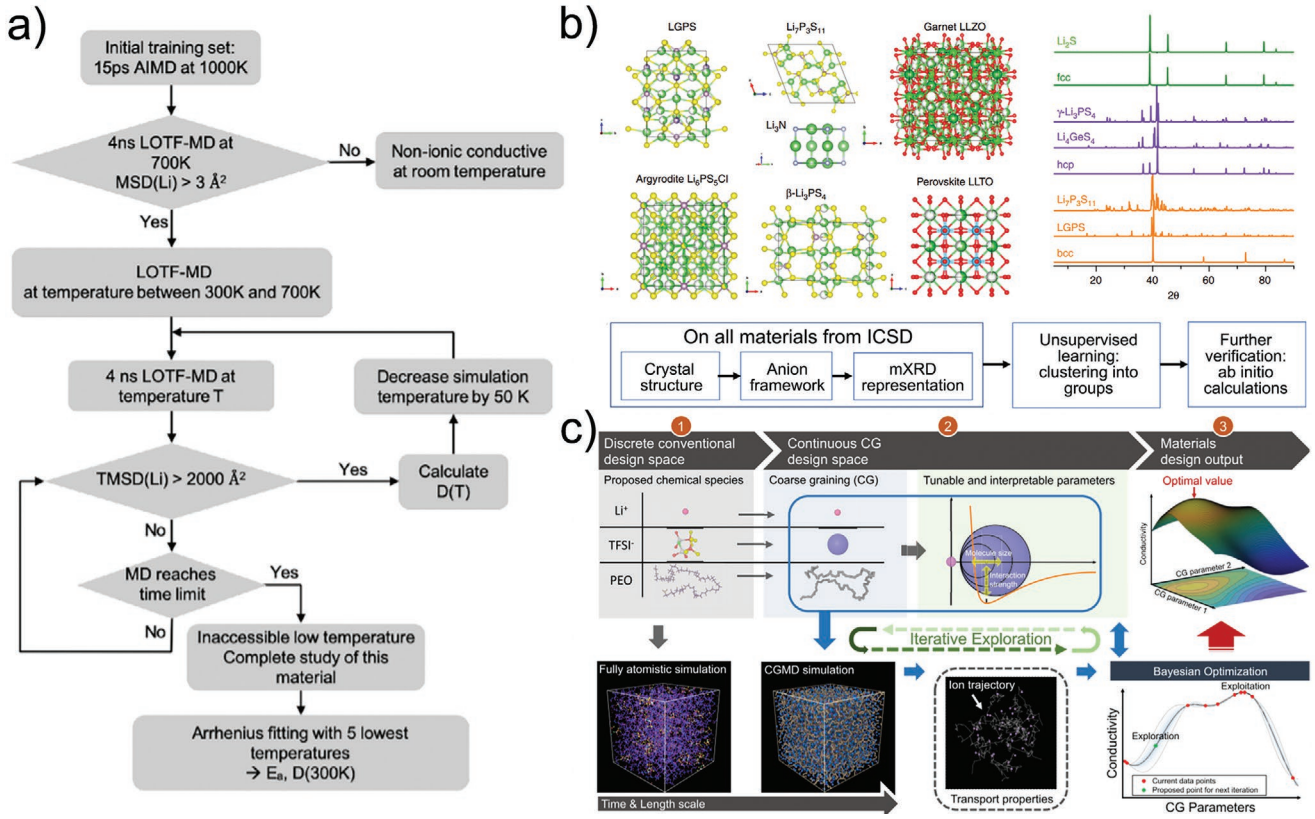


Figure 5. a) Flowchart of the "machine learning–molecular dynamics" ionic conductor screening process. b) Schematics of the unsupervised discovery of SSLCs. c) Illustration of the coarse-grained molecular dynamics–Bayesian optimization framework. a) Adapted with permission.[153] Copyright 2020, American Chemical Society. b) Adapted under the terms of the CC-BY Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0).[156] Copyright 2019, The Authors, published by Springer Nature. c) Adapted with permission.[157] Copyright 2020, American Chemical Society.

stability of the liquid electrolyte across the full operating range of LIBs. For the commercial liquid organic electrolytes, there are potential safety issues owing to their volatility and flammability. Thus, the design of novel solid electrolytes with high ionic conductivity is of great significance to develop all-solid-state LIBs since they are less flammable and safe. Zhang et al.[156] used the unsupervised learning approach to screen all known Li-containing compounds from ICSD. The trained model cluster the solid-state Li-ion conductors (SSLCs) into groups of compounds with high conductivity and other groups with poor ionic conduction. Conducting the ML model, 16 new fast Li-conductors with conductivities of 10^{-4} to 10^{-1} S cm^{-1} were discovered (Figure 5b). The discovered candidates have entirely different structures and chemical compositions compared with the current known fast Li-ion conductors. Recently, a new framework for designing solid polymer electrolytes was proposed by Wang et al,[157] which combined coarse-grained MD with Bayesian optimization (CGMD-BO), as shown in Figure 5c. The CG simulation could preserve molecular-level information, constructing a continuous high-dimensional design space. The BO algorithm showed the unique advantages of efficiency and flexibility in optimizing lithium ionic conductivity with the molecular-level material properties as descriptors. The CGMD-BO framework was expected to be an emerging approach for designing other complex multicomponent material systems.

Several aforementioned successful examples have manifested the enormous potential and advantage of the ML method in discovering new materials and revealing the structure–property relationship. However, there are still lots of challenges and a huge space for further development of ML in material design. Owing to the lack of datasets availability, some dilemma still exists in the data collection stage for machine learning methods. In addition, to evaluate the properties of the different structures, the available method is constructing ML models for quantitative structure property relationship or to convert the structure prediction into an optimization problem. Nevertheless, the current framework for structure prediction can only predict the known structures instead of unknown structures. Thus, it is urgent to address the issue that how to incorporate the domain knowledge summarized in experiments or implicit in heuristic rules for new-type structural prediction when developing ML models.

Given the challenges in model parameterization and the highly nonlinear and coupled nature of battery degradation processes, many efforts have been devoted to seeking high-efficiency data-driven approaches for the predication of materials and states. Many ML methods, such as ANN. SVM, DT, etc., have been commonly employed in previous reports. Nevertheless, a major disadvantage is that as-employed ML approaches demand large amount of experimental training data to create an

accurate model. In addition, the reliability of these approaches is beyond the experimental training set.[145]

4. Challenges of Deploying AI in Real-World Scenarios and an Integrated Framework

While it has been highly touted that AI or ML would advance and transform a lot of domains, including applications in LIB territory,[64,68,72,156,158] practical deployment of AI/ML algorithms into real-world scenarios still faces great challenges that are model and data. Learning-based approaches can extract complicated and nonlinear patterns in the training datasets and translate the meta-data to statistical models that are able to implement prediction, classification, optimization, or detection problem for specific applications. In the most accepted format of the supervised learning model, it entails a two-phase process, that is, training and inference. In the training phase, ML algorithms absorb the labeled datasets to learn parameter/weight values in the model. In the inference phase, the trained model takes the input of new data points to make an inference about the corresponding label. However, in real-world applications with physical systems, this process faces two inherent challenges, as follows:

a) Data scarcity. Learning-based statistical model, especially the deep learning with myriad tunable parameter requires a massive amount of training data to ensure model quality. In the real physical system in the battery area, however, the collection and/or access to large amounts of data remains challenging due to high cost, long delays, and concerns over compliance and safety; for instance, the probability of a battery cell catastrophically failing related to historical conditions (e.g., voltage, operating rates, temperature, mechanical shock, etc.), as well as unmeasurable manufacturing defects. The challenge remains in accumulating training data in rare events/conditions, where involving failure scenarios would enhance the accuracy of DMMs and finally realize predictions of failures/anomalies.[159] Many pioneers have published datasets of battery failure comprising hundreds of cells, such as from NASA,[160] but the datasets are far smaller than what are required to grab the failure scenarios. Furthermore, owing to the sensitivity to operating conditions (e.g., temperature, load-profile, etc.) and highly variable conditions, the real-world lifespan estimation of batteries has cell-to-cell variations. These variations are often hard to measure and cover due to manufacturing inconsistencies.[161] All these challenges in data acquisition have seriously affected the deployment of ML algorithms in real-world scenarios.

b) Cost-safety concerns. Material search, property prediction, design, and management of battery systems are critical problems in the LIB economy and safety. For instance, regulating energy systems inevitably leads to intrinsic safety risks, and Li-ion technology is not exempt from incidents. It is unfortunate that the LIB is often the victim of fierce cost cutting, and people try to cram more and more energy in the same volume.[162] Further, it is of interest to predict the property/performance of the battery system over long horizons for safety reasons and service-level agreement with customers. A warning coming too early or too late may turn out to be incorrect maintenance or lead to ignoring the imminent failure, which renders the increase of system risks and costs that are unacceptable.[163] Additionally, the high cost of materials for LIBs (e.g., negative/anode electrode materials, electrolytes, etc.) not only promises battery chemistries, but also puts forward higher requirements for battery design and accurate material exploration in order to reduce the cost of performance assessment and manufacturing.[85,164] ML-based approaches are supposed to take into account all uncertainties and unpredictable events/conditions in the energy systems. Due to historical reasons (e.g., limited data, safety compliance, etc.), however, the management/operation/design of energy systems still depend to a large extent on the decision-making of human experts,[165] rendering the industry with a risk-averse mindset. As a consequence, these critical nature demand novel approaches in adopting ML solutions into the battery-system economy.

To tackle these daunting challenges, we have proposed to integrate ML algorithms of different nature into a unified framework (**Figure 6**), pivoting around the digital twin, to promote advanced applications in the battery-system economy. The

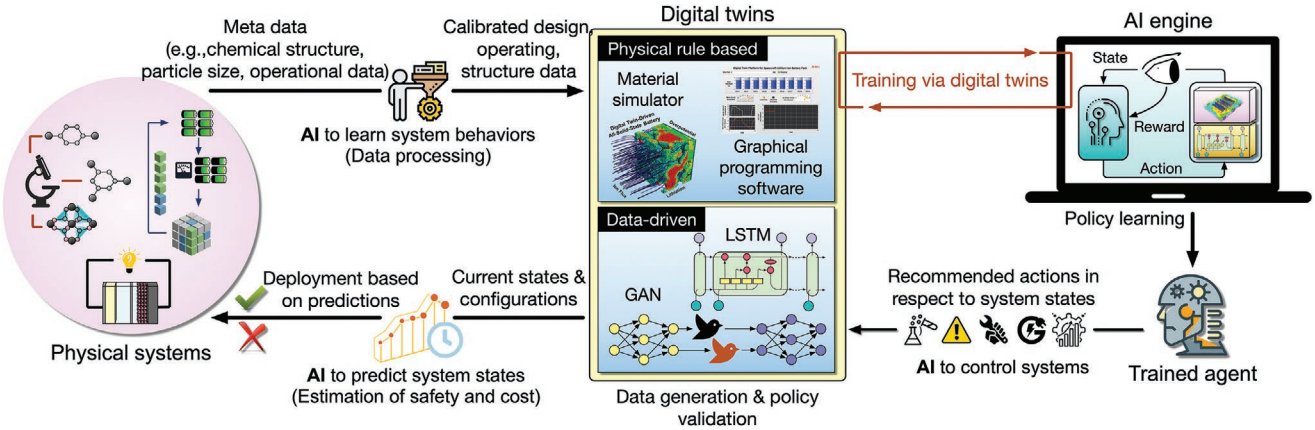


Figure 6. Illustration of a closed-loop framework for the design, material discovery, property prediction, and cell management of battery systems using digital twins and AI capabilities.

proposed framework is depicted in Figure 6, including three modules, which are physical systems/scenarios, digital twin, and AI engine. The physical system represents the real environment/system which is the target (e.g., battery design, material discovery, cell management, etc.) of learning-based approaches. The digital twin represents the digitalized cyber environments/models of the real-world scenarios. Physical and data-driven models are available to construct the digital twin, where the physical model is often constructed based on the physical rules, and the DMM is trained via massive historical data from the real environment or system. Both of them aim to precisely imitate the dynamics and behaviors of the real environment and are able to synthesize additional datasets and assess the performance of AI-based optimization approaches. AI engine represents the learning-based approaches to optimize/diagnose/control real-world systems. In the proposed framework, the AI engine is recommended to interact with the digital twin to avoid the risks of adopting ML algorithms on the physical system as well as reduce the verification costs of the experiment and trial production. These three modules interact among themselves through different forces (i.e., AI capabilities), as follows:

a) The route from the physical system to the digital twin represents the digitalization of the real-world environment via descriptive AI/ML. Raw data (e.g., chemical structure, particle size, etc.) is sampled from the physical system. The descriptive AI aims to ensure the data quality and analyze the composition of data to understand the complex behaviors of the system to model the digital twin. The digital twin has high flexibility and can be a set of formulations, simulators, or learning-based models.
b) The cyclic route between the digital twin and AI engine represents the implementation of prescriptive AI on the digital twin to optimize/solve the problems of the real-world systems. Instead of the physical system or component, these learning algorithms can directly interact with the digital twin to grab internal behaviors and learn complicated patterns of the system. The digital twin can synthesize the data under any condition in a relatively short time that enriches the diversity of the training dataset used by the AI engine. In addition, the route between the well-trained AI agent and the digital twin presents that the digital twin can conduct the validation before the deployment of AI recommended actions, resulting in cost savings and safety improvement.
c) The route from digital twin to physical system represents the prediction of the states in the future given specified inputs. The data-driven twin can also learn system behaviors from the historical data via the predictive AI. The sophisticated AI model is then able to predict future states (e.g., lifetime, SOC, SOH, etc.) of the system without exposing the physical system to the uncertainties of the AI algorithm. Furthermore, based on the predicted results, proper actions can also be implemented in the physical system to guarantee the stability or improve the performance of the system.

The proposed framework integrates the ML approaches with different nature (i.e., prescriptive, descriptive, and predictive ML/AL) and the three modules (i.e., physical system, digital twin, and AI engine) to offer the benefits as follows:

a) Data enrichment. Enriching the datasets to achieve high quality and diversity plays a critical role in ML algorithms. Descriptive AI captures inherent mechanisms of the target system to model accurate digital twin, which guarantees the synthetic data quality. Specifically, a well-constructed digital twin is able to synthesize diverse data (e.g., failures, anomalies, data under different conditions) that are rarely observed from the real-world environment or may put the system in jeopardy. The trustable data improve the diversity of training datasets and will then empower the AI engine and elevate the stability and applicability of ML approaches.
b) Efficient and safe deployment. The digital twin is the key component of the proposed framework and is used to conduct the pre-training and validate the effects of ML-based optimizations. In the proposed framework, training the AI engine on the digital twin in lieu of the physical system puts an end to the risks of misoperation and/or incorrect recommendations. Additionally, as the training data is finite, it is not ensured that the promising early results obtained in current training datasets will actually translate to correct deployment due to the engineering complexity. Consequently, the high-fidelity digital twin is then used to validate the recommended actions from the AI engine in advance, to evaluate the safety and estimate the performance.

5. Summary and Outlook

Computational chemistry has become a mature approach to complement and aid experimental studies for predicting and designing new materials, and many materials databases have been developed. The combination of AI with the materials databases is promising to accelerate materials innovation for batteries. Moreover, machine learning modes are also powerful in developing BMS, which is closely related to the health and safety of LIBs.

To promote electrochemical performance, it is vital to develop suitable electrode and electrolyte materials. Theoretically, novel materials can be discovered with the aid of predicting properties. However, in many instances, it is very challenging and expensive to determine the properties via large-scale experiments or DFT computations. ML algorithms can "learn" complex correlations and patterns from the existing data, providing a solution to the rapid screen of materials. The basic workflow for battery material property prediction via ML methods as follows: first, feature engineering is beneficial to identify the conditional attributes; second, the mapping relationship is established between these conditional factors and the decision attributes through model training; last, various properties (battery voltage, ionic conductivity, etc.) can be predicted by the trained model.

Although the screened battery materials may lead to better performance and more complicated battery dynamics, the safety of batteries could be another concern, especially for EV batteries. Accurate determination of the SOC/SOH and reliable prediction of RUL will alleviate the problem and unlock improvements in battery manufacturing, usage, and optimization. An intelligent BMS that can predict and monitor battery behavior is critical for both end-users and manufacturers.

Thus, predicting the degradation behavior of batteries by ML techniques is also essential for the entire electrification system. The fundamental goal of ML models in rechargeable batteries is to establish the QSAR between conditional attributes and decision attributes through low-cost and accurate predictions.

Modern computational tools can predict the properties of particular materials under specific conditions. However, up to now, the inverse design of battery materials has been computationally infeasible due to the massive complexity. Contrary to the property prediction, the properties of materials are the input, and the structure and composition are the output in an inverse materials design. The key issue is to identify promising chemical components and structures of materials, which can be synthesized in the lab. Aiming at the discovery and design of materials, the first step is to generate the key descriptors or features that are closely associated with the material properties of interest. The next step is to construct an accurate model between the descriptors and target properties. Theoretically, according to the ML model trained by a given dataset (materials → properties), the inverse design can be conducted to discover new materials with the intended properties.

While it has been highly touted that AI or ML would advance and transform the LIB territory, practical deployment of AI/ML algorithms into real-world scenarios still faces great challenges. Learning-based approaches can extract complicated and nonlinear patterns from the training datasets and translate the metadata to statistical models. In the most accepted format of the supervised learning model, it entails a two-phase process, that is, training and inference. However, in real-world applications with physical systems, this process faces two inherent challenges, that is, data scarcity and cost safety concern. To tackle these challenges, we propose to integrate ML algorithms of different nature into a unified framework, pivoting around the digital twin, to promote advanced applications in the battery-system economy. The proposed framework includes three modules, which are physical systems/scenarios, digital twin, and AI engine. These three modules interact among themselves through different forces (i.e., AI capabilities). The proposed framework integrates the ML approaches with different nature (i.e., prescriptive, descriptive, and predictive ML/AL) and the three modules to offer the benefits, such as data enrichment, efficiency, and safe deployment.

Acknowledgements

C.L., X.Z., and L.Z. contributed equally to this work. Q.Y. acknowledges the funding support from Singapore MOE AcRF Tier 1 grant nos. 2020-T1-001-031, and Tier 2 grant nos. 2017-T2-2-069. Y.W. acknowledges the Nation Research Foundation, Prime Minister's Office, Singapore under its Energy Programme (EP Award No. NRF2017EWT-EP003-023) administrated by the Energy Market Authority of Singapore; its Green Data Centre Research (GDCR Award No. NRF2015ENC-GDCR01001-003) administrated by the Info-communications Media Development Authority. M.S. gratefully acknowledges the financial support from National Research foundation of Singapore Investigatorship Award Number NRFI2017-08 and AStar AME programmatic funding number A20H3g2140. S.L. acknowledges the financial support from the Academic Research Fund Tier 1 (RG8/20), Tier 1 (RG104/18) and the computing resources from National Supercomputing Centre Singapore. The authors also like to acknowledge 111 project (D18023) from Zhengzhou University for their support for this work.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

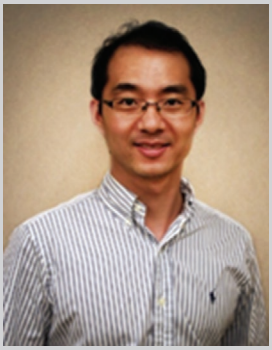
lithium-ion batteries, machine learning, materials discovery and prediction, state prediction

Received: February 22, 2021
Revised: May 24, 2021
Published online:

[1] D. Lindley, *Nature* **2010**, *463*, 18.
[2] V. Etacheri, R. Marom, R. Elazari, G. Salitra, D. Aurbach, *Energy Environ. Sci.* **2011**, *4*, 3243.
[3] A. Manthiram, X. Yu, S. Wang, *Nat. Rev. Mater.* **2017**, *2*, 16103.
[4] Y. S. Meng, M. E. Arroyo-de Dompablo, *Energy Environ. Sci.* **2009**, *2*, 589.
[5] A. Ullah, A. Majid, N. Rani, *J. Energy Chem.* **2018**, *27*, 219.
[6] H. Heinz, H. Ramezani-Dakhel, *Chem. Soc. Rev.* **2016**, *45*, 412.
[7] M. Endo, C. Kim, K. Nishimura, T. Fujino, K. Miyashita, *Carbon* **2000**, *38*, 183.
[8] R. N. Methekar, P. W. C. Northrop, K. Chen, R. D. Braatz, V. R. Subramanian, *J. Electrochem. Soc.* **2011**, *158*, A363.
[9] A. Belsky, M. Hellenbrandt, V. L. Karen, P. Luksch, *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 364.
[10] C. R. Groom, I. J. Bruno, M. P. Lightfoot, S. C. Ward, *Acta Crystallogr., Sect. B: Struct. Sci.* **2016**, *72*, 171.
[11] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Mater.* **2013**, *1*, 011002.
[12] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. W. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, O. Levy, *Comput. Mater. Sci.* **2012**, *58*, 227.
[13] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, C. Wolverton, *JOM* **2013**, *65*, 1501.
[14] J. Hachmann, R. Olivares-Amaya, S. Atahan-Evrenk, C. Amador-Bedolla, R. S. Sánchez-Carrera, A. Gold-Parker, L. Vogt, A. M. Brockway, A. Aspuru-Guzik, *J. Phys. Chem. Lett.* **2011**, *2*, 2241.
[15] C. Ortiz, O. Eriksson, M. Klintenberg, *Comput. Mater. Sci.* **2009**, *44*, 1042.
[16] J. Jie, M. Weng, S. Li, D. Chen, S. Li, W. Xiao, J. Zheng, F. Pan, L. Wang, *Sci. China: Technol. Sci.* **2019**, *62*, 1423.
[17] C. Chen, Y. Zuo, W. Ye, X. Li, Z. Deng, S. P. Ong, *Adv. Energy Mater.* **2020**, *10*, 1903242.
[18] W. Sha, Y. Guo, Q. Yuan, S. Tang, X. Zhang, S. Lu, X. Guo, Y.-C. Cao, S. Cheng, *Adv. Intell. Syst.* **2020**, *2*, 1900143.
[19] T. Zhou, Z. Song, K. Sundmacher, *Engineering* **2019**, *5*, 1017.
[20] K. K. Yang, Z. Wu, F. H. Arnold, *Nat. Methods* **2019**, *16*, 687.
[21] J. Wei, X. Chu, X.-Y. Sun, K. Xu, H.-X. Deng, J. Chen, Z. Wei, M. Lei, *InfoMat* **2019**, *1*, 338.
[22] R. Jose, S. Ramakrishna, *Appl. Mater. Today* **2018**, *10*, 127.
[23] A. Agrawal, A. Choudhary, *APL Mater.* **2016**, *4*, 053208.
[24] A. Chen, X. Zhang, Z. Zhou, *InfoMat* **2020**, *2*, 553.
[25] G. H. Gu, J. Noh, I. Kim, Y. Jung, *J. Mater. Chem. A* **2019**, *7*, 17096.
[26] R. Jalem, M. Nakayama, T. Kasuga, *J. Mater. Chem. A* **2014**, *2*, 720.
[27] A. D. Sendek, Q. Yang, E. D. Cubuk, K.-A. N. Duerloo, Y. Cui, E. J. Reed, *Energy Environ. Sci.* **2017**, *10*, 306.
[28] Z. Ahmad, T. Xie, C. Maheshwari, J. C. Grossman, V. Viswanathan, *ACS Cent. Sci.* **2018**, *4*, 996.

[29] A. Bhowmik, I. E. Castelli, J. M. Garcia-Lastra, P. B. Jørgensen, O. Winther, T. Vegge, *Energy Storage Mater.* **2019**, *21*, 446.
[30] Y. Liu, B. Guo, X. Zou, Y. Li, S. Shi, *Energy Storage Mater.* **2020**, *31*, 434.
[31] H. Xu, J. Zhu, D. P. Finegan, H. Zhao, X. Lu, W. Li, N. Hoffman, A. Bertei, P. Shearing, M. Z. Bazant, *Adv. Energy Mater.* **2021**, *11*, 2003908.
[32] R. E. Brandt, R. C. Kurchin, V. Steinmann, D. Kitchaev, C. Roat, S. Levcenco, G. Ceder, T. Unold, T. Buonassisi, *Joule* **2017**, *1*, 843.
[33] A. Mistry, A. A. Franco, S. J. Cooper, S. A. Roberts, V. Viswanathan, *ACS Energy Lett.* **2021**, *6*, 1422.
[34] L. Lu, X. Han, J. Li, J. Hua, M. Ouyang, *J. Power Sources* **2013**, *226*, 272.
[35] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, P. Van den Bossche, *Renewable Sustainable Energy Rev.* **2016**, *56*, 572.
[36] X. Feng, M. Ouyang, X. Liu, L. Lu, Y. Xia, X. He, *Energy Storage Mater.* **2018**, *10*, 246.
[37] M. Shen, Q. Gao, *Int. J. Energy Res.* **2019**, *43*, 5042.
[38] K. Liu, K. Li, Q. Peng, C. Zhang, *Front. Mech. Eng.* **2019**, *14*, 47.
[39] Y. Li, K. Liu, A. M. Foley, A. Zülke, M. Berecibar, E. Nanini-Maury, J. Van Mierlo, H. E. Hoster, *Renewable Sustainable Energy Rev.* **2019**, *113*, 109254.
[40] D. N. T. How, M. A. Hannan, M. S. H. Lipu, P. J. Ker, *IEEE Access* **2019**, *7*, 136116.
[41] M.-F. Ng, J. Zhao, Q. Yan, G. J. Conduit, Z. W. Seh, *Nat. Mach. Intell.* **2020**, *2*, 161.
[42] C. Vidal, P. Malysz, P. Kollmeyer, A. Emadi, *IEEE Access* **2020**, *8*, 52796.
[43] A. Jokar, B. Rajabloo, M. Désilets, M. Lacroix, *J. Power Sources* **2016**, *327*, 44.
[44] K. Doya, *Neural Networks* **1999**, *12*, 961.
[45] P. Chemouil, P. Hui, W. Kellerer, Y. Li, R. Stadler, D. Tao, Y. Wen, Y. Zhang, *IEEE J. Sel. Areas Commun.* **2019**, *37*, 1185.
[46] T. H. Davenport, *Harv. Bus. Rev.* **2013**, *91*, 64.
[47] M. Aykol, P. Herring, A. Anapolsky, *Nat. Rev. Mater.* **2020**, *5*, 725.
[48] J. J. Irwin, B. K. Shoichet, *J. Chem. Inf. Model.* **2005**, *45*, 177.
[49] S. Curtarolo, D. Morgan, K. Persson, J. Rodgers, G. Ceder, *Phys. Rev. Lett.* **2003**, *91*, 135503.
[50] R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S. I. Hong, M. Baldo, R. P. Adams, A. Aspuru-Guzik, *Nat. Mater.* **2016**, *15*, 1120.
[51] G. L. Guimaraes, B. Sanchez-Lengeling, C. Outeiral, P. L. C. Farias, A. Aspuru-Guzik, *arXiv: 1705.10843,* **2017**.
[52] S. Kench, S. J. Cooper, *Nat. Mach. Intell.* **2021**, *3*, 299.
[53] B. Sanchez-Lengeling, C. Outeiral, G. L. Guimaraes, A. Aspuru-Guzik, *ChemRxiv* **2017**, https://doi.org/10.26434/chemrxiv.5309668.v3.
[54] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, R. D. Braatz, *Nat. Energy* **2019**, *4*, 383.
[55] J. Li, X. Li, D. He, *IEEE Access* **2019**, *7*, 75464.
[56] S. Zhu, C. He, N. Zhao, J. Sha, *J. Power Sources* **2021**, *482*, 228983.
[57] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, A. A. Lee, *Nat. Commun.* **2020**, *11*, 1706.
[58] N. Harting, R. Schenkendorf, N. Wolff, U. Krewer, *Appl. Sci.* **2018**, *8*, 821.
[59] A. Naha, A. Khandelwal, S. Agarwal, P. Tagade, K. S. Hariharan, A. Kaushik, A. Yadu, S. M. Kolake, S. Han, B. Oh, *Sci. Rep.* **2020**, *10*, 1301.
[60] B. Wu, S. Han, K. G. Shin, W. Lu, *J. Power Sources* **2018**, *395*, 128.
[61] M. Attarian Shandiz, R. Gauvin, *Comput. Mater. Sci.* **2016**, *117*, 270.
[62] X. Wang, R. Xiao, H. Li, L. Chen, *J. Materiomics* **2017**, *3*, 178.
[63] T. Sarkar, A. Sharma, A. K. Das, D. Deodhare, M. D. Bharadwaj, in *2014 2nd Int. Conf. on Devices, Circuits and Systems (ICDCS)*, IEEE, Piscataway, NJ, USA **2014**, https://doi.org/10.1109/ICDCSyst.2014.6926140.
[64] R. P. Joshi, J. Eickholt, L. Li, M. Fornari, V. Barone, J. E. Peralta, *ACS Appl. Mater. Interfaces* **2019**, *11*, 18494.
[65] R. A. Eremin, P. N. Zolotarev, O. Y. Ivanshina, I. A. Bobrikov, *J. Phys. Chem. C* **2017**, *121*, 28293.
[66] R. P. Cunha, T. Lombardo, E. N. Primo, A. A. Franco, *Batteries Supercaps* **2020**, *3*, 60.
[67] P. Hennig, M. Kiefel, *J. Mach. Learn. Res.* **2013**, *14*, 843.
[68] O. Allam, B. W. Cho, K. C. Kim, S. S. Jang, *RSC Adv.* **2018**, *8*, 39414.
[69] T. Parthiban, R. Ravi, N. Kalaiselvi, *Electrochim. Acta* **2007**, *53*, 1877.
[70] K. Sodeyama, Y. Igarashi, T. Nakayama, Y. Tateyama, M. Okada, *Phys. Chem. Chem. Phys.* **2018**, *20*, 22585.
[71] A. Ishikawa, K. Sodeyama, Y. Igarashi, T. Nakayama, Y. Tateyama, M. Okada, *Phys. Chem. Chem. Phys.* **2019**, *21*, 26399.
[72] Y. Okamoto, Y. Kubo, *ACS Omega* **2018**, *3*, 7868.
[73] L. D. Ellis, S. Buteau, S. G. Hames, L. M. Thompson, D. S. Hall, J. R. Dahn, *J. Electrochem. Soc.* **2018**, *165*, A256.
[74] A. M. Nolan, Y. Zhu, X. He, Q. Bai, Y. Mo, *Joule* **2018**, *2*, 2016.
[75] R. Jalem, T. Aoyama, M. Nakayama, M. Nogami, *Chem. Mater.* **2012**, *24*, 1357.
[76] R. Jalem, M. Kimura, M. Nakayama, T. Kasuga, *J. Chem. Inf. Model.* **2015**, *55*, 1158.
[77] R. Jalem, K. Kanamori, I. Takeuchi, M. Nakayama, H. Yamasaki, T. Saito, *Sci. Rep.* **2018**, *8*, 5845.
[78] S. Ibrahim, M. R. Johan, *Int. J. Electrochem. Sci.* **2011**, *6*, 5565.
[79] K. Fujimura, A. Seko, Y. Koyama, A. Kuwabara, I. Kishida, K. Shitara, C. A. J. Fisher, H. Moriwake, I. Tanaka, *Adv. Energy Mater.* **2013**, *3*, 980.
[80] N. Kireeva, V. S. Pervov, *Phys. Chem. Chem. Phys.* **2017**, *19*, 20904.
[81] K. Miwa, R. Asahi, *Phys. Rev. Mater.* **2018**, *2*, 105404.
[82] B. Liu, J. Yang, H. Yang, C. Ye, Y. Mao, J. Wang, S. Shi, J. Yang, W. Zhang, *J. Mater. Chem. A* **2019**, *7*, 19961.
[83] M. B. Dixit, A. Verma, W. Zaman, X. Zhong, P. Kenesei, J. S. Park, J. Almer, P. P. Mukherjee, K. B. Hatzell, *ACS Appl. Energy Mater.* **2020**, *3*, 9534.
[84] J. M. Tarascon, M. Armand, *Nature* **2001**, *414*, 359.
[85] R. Schmuch, R. Wagner, G. Hörpel, T. Placke, M. Winter, *Nat. Energy* **2018**, *3*, 267.
[86] A. Kwade, W. Haselrieder, R. Leithoff, A. Modlinger, F. Dietrich, K. Droeder, *Nat. Energy* **2018**, *3*, 290.
[87] S. B. Peterson, J. Apt, J. F. Whitacre, *J. Power Sources* **2010**, *195*, 2385.
[88] V. Ramadesigan, P. W. C. Northrop, S. De, S. Santhanagopalan, R. D. Braatz, V. R. Subramanian, *J. Electrochem. Soc.* **2012**, *159*, R31.
[89] W. Waag, C. Fleischer, D. U. Sauer, *J. Power Sources* **2014**, *258*, 321.
[90] X. Hu, S. E. Li, Y. Yang, *IEEE Trans. Transp. Electrif.* **2016**, *2*, 140.
[91] M. Berecibar, F. Devriendt, M. Dubarry, I. Villarreal, N. Omar, W. Verbeke, J. Van Mierlo, *J. Power Sources* **2016**, *320*, 239.
[92] Y. Zhang, R. Xiong, H. He, Z. Liu, in *2017 Prognostics and System Health Management Conf. (PHM-Harbin)*, IEEE, Piscataway, NJ, USA **2017**, https://doi.org/10.1109/PHM.2017.8079316.
[93] T. Wu, M. Wang, Q. Xiao, X. Wang, *Smart Grid Renewable Energy* **2012**, *3*, 51.
[94] K.-H. Tseng, J.-W. Liang, W. Chang, S.-C. Huang, *Energies* **2015**, *8*, 2889.
[95] N. Dawson-Elli, S. B. Lee, M. Pathak, K. Mitra, V. R. Subramanian, *J. Electrochem. Soc.* **2018**, *165*, A1.
[96] P. Khumprom, N. Yodo, *Energies* **2019**, *12*, 660.
[97] J. Wu, Y. Wang, X. Zhang, Z. Chen, *J. Power Sources* **2016**, *327*, 457.
[98] D. P. Finegan, J. Zhu, X. Feng, M. Keyser, M. Ulmefors, W. Li, M. Z. Bazant, S. J. Cooper, *Joule* **2021**, *5*, 316.

[99] M. Schimpe, M. E. von Kuepach, M. Naumann, H. C. Hesse, K. Smith, A. Jossen, *J. Electrochem. Soc.* **2018**, *165*, A181.
[100] Y. Wang, J. Tian, Z. Sun, L. Wang, R. Xu, M. Li, Z. Chen, *Renewable Sustainable Energy Rev.* **2020**, *131*, 110015.
[101] W. He, M. Pecht, D. Flynn, F. Dinmohammadi, *Energies* **2018**, *11*, 2120.
[102] M. Aykol, C. B. Gopal, A. Anapolsky, P. K. Herring, B. van Vlijmen, M. D. Berliner, M. Z. Bazant, R. D. Braatz, W. C. Chueh, B. D. Storey, *J. Electrochem. Soc.* **2021**, *168*, 030525.
[103] P. Gasper, K. Gering, E. Dufek, K. Smith, *J. Electrochem. Soc.* **2021**, *168*, 020502.
[104] S. R. Hashemi, A. Bahadoran Baghbadorani, R. Esmaeeli, A. Mahajan, S. Farhad, *Int. J. Energy Res.* **2021**, *45*, 5747.
[105] Y. Li, M. Abdel-Monem, R. Gopalakrishnan, M. Berecibar, E. Nanini-Maury, N. Omar, P. van den Bossche, J. Van Mierlo, *J. Power Sources* **2018**, *373*, 40.
[106] L. Kang, X. Zhao, J. Ma, *Appl. Energy* **2014**, *121*, 20.
[107] S. Tong, J. H. Lacap, J. W. Park, *J. Energy Storage* **2016**, *7*, 236.
[108] E. Chemali, P. J. Kollmeyer, M. Preindl, A. Emadi, *J. Power Sources* **2018**, *400*, 242.
[109] T. H. Donato, M. G. Quiles, *Adv. Comput. Intell.* **2018**, *5*, 1.
[110] D. Jiménez-Bermejo, J. Fraile-Ardanuy, S. Castaño-Solis, J. Merino, R. Álvaro-Hermana, *Procedia Comput. Sci.* **2018**, *130*, 533.
[111] A. G. Kashkooli, H. Fathiannasab, Z. Mao, Z. Chen, *J. Electrochem. Soc.* **2019**, *166*, A605.
[112] J. Kleiner, M. Stuckenberger, L. Komsiyska, C. Endisch, *Batteries* **2021**, *7*, 31.
[113] V. Chandran, C. K. Patil, A. Karthick, D. Ganeshaperumal, R. Rahim, A. Ghosh, *World Electr. Veh. J.* **2021**, *12*, 38.
[114] J. C. Álvarez Antón, P. J. García Nieto, F. J. de Cos Juez, F. Sánchez Lasheras, M. González Vega, M. N. Roqueñí Gutiérrez, *Appl. Math. Model.* **2013**, *37*, 6244.
[115] J. N. Hu, J. J. Hu, H. B. Lin, X. P. Li, C. L. Jiang, X. H. Qiu, W. S. Li, *J. Power Sources* **2014**, *269*, 682.
[116] L. Yao, Z. Fang, Y. Xiao, J. Hou, Z. Fu, *Energy* **2021**, *214*, 118866.
[117] Y. Jia, J. Li, C. Yuan, X. Gao, W. Yao, M. Lee, J. Xu, *Adv. Energy Mater.* **2021**, *11*, 2003868.
[118] T. Lan, K. Jermsittiparsert, S. T. Alrashood, M. Rezaei, L. Al-Ghussain, M. A. Mohamed, *Energies* **2021**, *14*, 569.
[119] Z. Fei, F. Yang, K.-L. Tsui, L. Li, Z. Zhang, *Energy* **2021**, *225*, 120205.
[120] T. Sun, R. Wu, Y. Cui, Y. Zheng, *J. Energy Storage* **2021**, *39*, 102594.
[121] C. Li, C. Zhang, in *2020 IEEE Int. Conf. on Prognostics and Health Management (ICPHM)*, IEEE, Piscataway, NJ, USA **2020**, https://doi.org/10.1109/ICPHM49022.2020.9187060.
[122] Y. Zhang, Z. Peng, Y. Guan, L. Wu, *Energy* **2021**, *221*, 119901.
[123] D. Roman, S. Saxena, V. Robu, M. Pecht, D. Flynn, *Nat. Mach. Intell.* **2021**, *3*, 447.
[124] A. A. Hussein, *Int. J. Mod. Nonlinear Theor. Appl.* **2014**, *3*, 199.
[125] C. Huang, Z. Wang, Z. Zhao, L. Wang, C. S. Lai, D. Wang, *IEEE Access* **2018**, *6*, 27617.
[126] X. Tan, D. Zhan, P. Lyu, J. Rao, Y. Fan, *J. Power Sources* **2021**, *484*, 229233.
[127] J. Chen, X. Feng, L. Jiang, Q. Zhu, *Energy* **2021**, *227*, 120451.
[128] Y. Yang, *Appl. Energy* **2021**, *292*, 116897.
[129] K. Kaur, A. Garg, X. Cui, S. Singh, B. K. Panigrahi, *Int. J. Energy Res.* **2021**, *45*, 3113.
[130] T. Nguyen-Thoi, X. Cui, A. Garg, L. Gao, T. T. Truong, *Energy Technol.* **2021**, *9*, 2100048.
[131] Y. Guo, Z. Yang, K. Liu, Y. Zhang, W. Feng, *Energy* **2021**, *219*, 119529.
[132] L. Chen, H. Wang, B. Liu, Y. Wang, Y. Ding, H. Pan, *Energy* **2021**, *215*, 119078.
[133] C. Hu, G. Jain, C. Schmidt, C. Strief, M. Sullivan, *J. Power Sources* **2015**, *289*, 105.
[134] G. O. Sahinoglu, M. Pajovic, Z. Sahinoglu, Y. Wang, P. V. Orlik, T. Wada, *IEEE Trans. Ind. Electron.* **2018**, *65*, 4311.
[135] E. Kwak, S. Jeong, J.-h. Kim, K.-Y. Oh, *J. Power Sources* **2021**, *483*, 229079.
[136] W. Li, J. Zhu, Y. Xia, M. B. Gorji, T. Wierzbicki, *Joule* **2019**, *3*, 2703.
[137] P. Fermín-Cueto, E. McTurk, M. Allerhand, E. Medina-Lopez, M. F. Anjos, J. Sylvester, G. dos Reis, *Energy AI* **2020**, *1*, 100006.
[138] S. S. Mansouri, P. Karvelis, G. Georgoulas, G. Nikolakopoulos, *IFAC-PapersOnLine* **2017**, *50*, 4727.
[139] A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, K. Dietmayer, *J. Power Sources* **2013**, *239*, 680.
[140] D. Yang, Y. Wang, R. Pan, R. Chen, Z. Chen, *Energy Procedia* **2017**, *105*, 2059.
[141] T. Zahid, K. Xu, W. Li, C. Li, H. Li, *Energy* **2018**, *162*, 871.
[142] J. Guo, Z. Li, M. Pecht, *J. Power Sources* **2015**, *281*, 173.
[143] L. Ren, L. Zhao, S. Hong, S. Zhao, H. Wang, L. Zhang, *IEEE Access* **2018**, *6*, 50587.
[144] R. R. Richardson, M. A. Osborne, D. A. Howey, *J. Power Sources* **2017**, *357*, 209.
[145] B. Wu, W. D. Widanage, S. Yang, X. Liu, *Energy AI* **2020**, *1*, 100016.
[146] A. Jain, J. A. Bollinger, T. M. Truskett, *AIChE J.* **2014**, *60*, 2732.
[147] B. Sanchez-Lengeling, A. Aspuru-Guzik, *Science* **2018**, *361*, 360.
[148] R. Olivares-Amaya, C. Amador-Bedolla, J. Hachmann, S. Atahan-Evrenk, R. S. Sánchez-Carrera, L. Vogt, A. Aspuru-Guzik, *Energy Environ. Sci.* **2011**, *4*, 4849.
[149] M.-H. Lin, J.-F. Tsai, C.-S. Yu, *Math. Probl. Eng.* **2012**, *2012*, 756023.
[150] J. C. Spall, *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*, Series in Discrete Mathematics and Optimization, John Wiley & Sons, Hoboken, NJ, USA **2003**.
[151] K. Alberi, M. B. Nardelli, A. Zakutayev, L. Mitas, S. Curtarolo, A. Jain, M. Fornari, N. Marzari, I. Takeuchi, M. L. Green, M. Kanatzidis, M. F. Toney, S. Butenko, B. Meredig, S. Lany, U. Kattner, A. Davydov, E. S. Toberer, V. Stevanovic, A. Walsh, N.-G. Park, A. Aspuru-Guzik, D. P. Tabor, J. Nelson, J. Murphy, A. Setlur, J. Gregoire, H. Li, R. Xiao, A. Ludwig, L. W. Martin, A. M. Rappe, S.-H. Wei, J. Perkins, *J. Phys. D: Appl. Phys.* **2018**, *52*, 013001.
[152] G. Houchins, V. Viswanathan, *J. Chem. Phys.* **2020**, *153*, 054124.
[153] C. Wang, K. Aoyagi, P. Wisesa, T. Mueller, *Chem. Mater.* **2020**, *32*, 3741.
[154] K. Xu, *Chem. Rev.* **2004**, *104*, 4303.
[155] P. M. Tagade, S. P. Adiga, S. Pandian, M. S. Park, K. S. Hariharan, S. M. Kolake, *npj Comput. Mater.* **2019**, *5*, 127.
[156] Y. Zhang, X. He, Z. Chen, Q. Bai, A. M. Nolan, C. A. Roberts, D. Banerjee, T. Matsunaga, Y. Mo, C. Ling, *Nat. Commun.* **2019**, *10*, 5260.
[157] Y. Wang, T. Xie, A. France-Lanord, A. Berkley, J. A. Johnson, Y. Shao-Horn, J. C. Grossman, *Chem. Mater.* **2020**, *32*, 4144.
[158] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, F. Herrera, *Inf. Fusion* **2020**, *58*, 82.
[159] D. P. Finegan, S. J. Cooper, *Joule* **2019**, *3*, 2599.
[160] W. Q. Walker, J. J. Darst, D. P. Finegan, G. A. Bayles, K. L. Johnson, E. C. Darcy, S. L. Rickman, *J. Power Sources* **2019**, *415*, 207.
[161] K. Rumpf, M. Naumann, A. Jossen, *J. Energy Storage* **2017**, *14*, 224.
[162] J.-M. Tarascon, *Philos. Trans. R. Soc., A* **2010**, *368*, 3227.
[163] H. Wenzl, I. Baring-Gould, R. Kaiser, B. Y. Liaw, P. Lundsager, J. Manwell, A. Ruddell, V. Svoboda, *J. Power Sources* **2005**, *144*, 373.
[164] L. Gaines, R. Cuenca, *Costs of Lithium-Ion Batteries for Vehicles*, Argonne National Laboratory, Lemont, IL, USA **2000**.
[165] L. Wei, X. Xu, Gurudayal, J. Bullock, J. W. Ager, *Chem. Mater.* **2019**, *31*, 7340.



Shuzhou Li received his B.Sc. degree, M.Sc. degree, and Ph.D. degree from Nankai University, Peking University, and University of Wisconsin, Madison, USA, respectively. He had worked as a postdoctoral fellow at Northwestern University, USA. Then, he joined the School of Materials Science and Engineering, Nanyang Technological University (NTU) and is currently an associate professor. His research interests focus on exploring the chemical and physical properties of materials by theoretical and computational tools.

Yonggang Wen is Professor of Computer Science and Engineering at Nanyang Technological University (NTU), Singapore. He has also served as the Associate Dean (Research) at College of Engineering at NTU Singapore since 2018. He received his Ph.D. degree in electrical engineering and computer science from Massachusetts Institute of Technology (MIT), Cambridge, USA, in 2008. He serves on editorial boards for IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Multimedia, etc. His research interests include cloud computing, green data centers, distributed machine learning, blockchain, multimedia networks, and mobile computing. He is a Fellow of IEEE.

Qingyu Yan is currently a professor in the School of Materials Science and Engineering at Nanyang Technology University. He joined the School of Materials Science and Engineering of Nanyang Technological University as an assistant professor in 2008 and became a Professor in 2018. He is currently the Chair of the Electrochemical Society, Singapore Section. He has been a fellow of Royal Society of Chemistry since 2018. He has published >300 papers (with a total citation of >27 000 and an *h*-index of 86) on the research area: battery development; thermoelectric materials and electrocalytic process for energy conversion.