Machine learning in energy storage materials

Zhong-Hui Shen1,2 | Han-Xing Liu1,2 | Yang Shen3 | Jia-Mian Hu4 | Long-Qing Chen5 | Ce-Wen Nan3

1State Key Laboratory of Advanced Technology for Materials Synthesis and Processing, Center of Smart Materials and Devices, Wuhan University of Technology, Wuhan, China
2International School of Materials Science and Engineering, Wuhan University of Technology, Wuhan, China
3State Key Lab of New Ceramics and Fine Processing, School of Materials Science and Engineering, Tsinghua University, Beijing, China
4Department of Materials Science and Engineering, University of Wisconsin-Madison, Madison, Wisconsin, USA
5Department of Materials Science and Engineering, The Pennsylvania State University, University Park, Pennsylvania, USA

Correspondence
Zhong-Hui Shen, State Key Laboratory of Advanced Technology for Materials Synthesis and Processing, Center of Smart Materials and Devices, Wuhan University of Technology, 430070 Wuhan, China.
Email: zhshen@whut.edu.cn

Ce-Wen Nan, State Key Lab of New Ceramics and Fine Processing, School of Materials Science and Engineering, Tsinghua University, 100084 Beijing, China.
Email: cwnan@mail.tsinghua.edu.cn

Abstract
With its extremely strong capability of data analysis, machine learning has shown versatile potential in the revolution of the materials research paradigm. Here, taking dielectric capacitors and lithium-ion batteries as two representative examples, we review substantial advances of machine learning in the research and development of energy storage materials. First, a thorough discussion of the machine learning framework in materials science is presented. Then, we summarize the applications of machine learning from three aspects, including discovering and designing novel materials, enriching theoretical simulations, and assisting experimentation and characterization. Finally, a brief outlook is highlighted to spark more insights on the innovative implementation of machine learning in materials science.

KEYWORDS
dielectric capacitor, energy storage, lithium-ion battery, machine learning

1 | INTRODUCTION

The foreseeable exhaustion of fossil fuels and consequent environmental deterioration has triggered burgeoning worldwide demands in developing sustainable energy alternatives.[1–3] Some renewable or clean sources, such as wind, hydro, and solar energy, have been regarded as promising solutions to generate electrical energy.[4,5] However, no matter what the source of energy is, it inevitably needs to face a key challenge on how to efficiently store fluctuating energy in miscellaneous applications ranging from large power grids to electric vehicles and various portable devices.[6,7] Thus, energy storage is a crucial step to determine the efficiency, stability, and reliability of an electricity supply system.[8]

Up to now, dielectric capacitors (DCs) and lithium-ion batteries (LIBs) are two leading electrical energy storage technologies, as shown in Figure 1A.[9–11] Benefiting from the physical storage mechanism via electric dipoles as illustrated in Figure 1B, DCs endow ultrahigh power density (on the order of megawatt), fast response time (up to ~us), high working voltage (~MV m⁻¹), and long...
cycle life (>10^5 cycles), and thus serve as important basic components in a wide range of power systems.\textsuperscript{[12–17]} However, the low energy density of DCs (<0.1 Wh kg\(^{-1}\) in most cases) limits more widespread applications.\textsuperscript{[18]}

While for state-of-the-art LIBs based on electrochemical storage mechanism as presented in Figure 1C, the gravimetric energy density is approaching 300 Wh kg\(^{-1}\), which enables the rapid expansion of hybrid or all-electric vehicles, portable electronic devices, and other stationary applications.\textsuperscript{[19–21]} However, the explosive growth of the market has put forward much higher requirements for energy storage density.\textsuperscript{[13,22,23]} Moreover, their charge–discharge rate, lifetime, recyclability, and safety have also become obstacles for further advanced applications.\textsuperscript{[24–27]} Therefore, although both DCs and LIBs have been very significantly improved in recent decades, they are not capable of catching up with the mushrooming demand of tomorrow’s energy storage and power supply systems in terms of performance, durability, safety, cost, recyclability, and so on. Therefore, addressing the above major challenges requires research and development (R&D) of energy storage materials at an unprecedented pace and scale.

Research paradigm revolution in materials science by the advances of machine learning (ML) has sparked promising potential in speeding up the R&D pace of energy storage materials.\textsuperscript{[28–32]} On the one hand, the rapid development of computer technology has been the major driver for the explosion of ML and other computational simulations. Ever since the first computer (ENIAC, Electronic Numerical Integrator and Computer) was invented, the capabilities of computer systems have been substantially improved, including speed, accuracy, reliability, storage, adaptability, and so forth. For example, compared to 5000 addition operations per second of ENIAC, the calculation speed of a modern supercomputer has reached up to the order of PFlops (10\(^{15}\) floating point operation per second). Furthermore, advances in data storage capability have also enabled us to efficiently deal with a ton of matrix multiplication when performing complex ML models. On the other hand, ML, as a radically new and potent method, is transforming the field of discovery and design of energy storage materials in recent years.\textsuperscript{[33,34]} It could not only be used to understand the composition–structure–property–processing–performance linkages by encoding the domain knowledge into ML models but also realize property prediction, new materials discovery, multiobjective performance optimization and inverse design, and so forth.\textsuperscript{[35–40]} More importantly, ML is playing an essential role in addressing some challenges that are too difficult or time-consuming for traditional physical modeling.\textsuperscript{[41–43]} It is in the above context that the
The applications of ML on the paradigm shift have generated substantial excitement.\textsuperscript{[44-47]} While significant progress has been made, there is still much to be done.\textsuperscript{[48]} In particular, to promote even greater advances of ML in materials science, it is imperative to strengthen the cross-fusion between materials and computer/physics/mathematics.\textsuperscript{[49,50]}

This review aims at providing a critical overview of ML-driven R&D in energy storage materials to show how advanced ML technologies are successfully used to address various issues. First, we present a fundamental ML workflow and this section is organized along six basic steps involved in this process, as shown in Figure 2. Key concepts, approaches, examples, and challenges in each step will be discussed. Then, taking DCs and LIBs as two representative examples, we highlight recent advancements of ML in the R&D of energy storage materials from three aspects: discovering and designing novel materials, enriching theoretical simulations, and assisting experimentation and characterization. Finally, we outline some perspectives on future challenges and opportunities in ML for energy storage materials.

\section*{2 | ML WORKFLOW}

ML, as an offspring of artificial intelligence, is ubiquitous in our modern world.\textsuperscript{[51,52]} It could tell us that systems can, if trained, identify patterns, learn from data, and make decisions with or without supervision. In materials science, it has been widely used in every part of the R&D cycle from materials composition design and technological process optimization to property prediction and performance evaluation under service conditions. In this section, we summarize the ML workflow with six basic steps, as shown in Figure 2. Up to now, there are many systematical reviews on the ML workflow,\textsuperscript{[29,53-55]} and thus we no longer discuss everything involved in each step in detail. Here, we do wish to present some core issues and thinking in every step that may be helpful to promote better applications of ML in energy storage materials.

\subsection*{2.1 | Goal}

Before performing an ML project, the first step is to clarify the goal of this study, which will fundamentally guide how to build this model. The goal strongly determines the technical choices for each step from data and algorithm to reliability and stability of this model, as illustrated in Figure 2. Based on the type of tasks, ML models could be classified into the following categories: classification, regression, clustering, dimensionality reduction, and so forth. For example, to address a set of problems where the output variable can take continuous values is usually considered as a regression task. While for a classification task, the model aims at basically categorizing a set of data into different classes. In materials science, it should be noted that the choice of ML model highly depends on the nature of the research problem. For instance, identifying whether a dielectric material is ferroelectric or paraelectric is a classification task. Establishing the analytical relationship between the effective electrical conductivity of the composite electrolyte and the intrinsic parameters of each phase component belongs to a regression problem. In addition, for feature selection and extraction from complex and diverse data, the dimensionality reduction model is an appropriate choice to analyze inputs and reduce them to only relevant ones for the learning goal.\textsuperscript{[56]} In turn, identifying the goal of the ML model also needs to take into account the volume and quality of data and available features, such as crystal structure, chemical composition,
and microstructure. In essence, the goal is the commander of the whole ML workflow, and all steps in ML interact with each other. The details of data resources will be further discussed in the following Section 2.2. On the basis of the above different goals, there are a large number of algorithms to choose from, which will be discussed in Section 2.4.

2.2 | Data

Data, as one of the most cherished resources in the information age, are an essential prerequisite for ML. The sufficient quality and quantity of data strongly determine the reliability and accuracy of ML results.[57] However, data in materials science are scarce and discrete.[58,59] Thus, how to obtain high quality and sufficient quantity of materials data has been recognized as the core challenge in the applications of ML in materials science. Generally, data in materials science could be acquired from open-source databases and literature, or generated by high-throughput calculations and experiments.[35,44,53,57] Over the past few decades, a large number of materials database, such as Materials Project, AFLOW, OQMD, and so forth, have been enhanced and open to the public.[33,45] For a more comprehensive review of the available materials database, please refer to some excellent review articles.[53,60,61] Those available databases have become important data sources for ML. However, limited by the diversity and volume of existing databases, it is still difficult to meet different research needs, for example, lack of particular properties or poor coverage or low consistency. Here lies the dilemma that what kind of problems ML can solve highly depends on the available databases, rather than our actual needs or interests, which strongly stifles the enormous potential of ML.

To expand materials data, especially for experimental data, mining data resources from open-source literature has gained extensive interest.[62–64] However, the complexity and diversity of literature make automatic data mining technology face huge challenges. For example, intelligently mining the numerical values of polarization and electric field from different diagrams of the ferroelectric hysteresis loop is extremely hard due to the diversity of image format, location, style, size, and coordinatometer in different kinds of literature. Furthermore, performing high-throughput calculations or experiments to generate desired data is becoming an efficient and popular method. On the one hand, the fast proliferation of data by high-throughput calculations and experiments is an important source for many databases to generate and expand data.[65–71] For example, Materials Projects, one of the most popular materials database, contains ~10^5 crystal structures generated by density-functional theory calculations. On the other hand, according to the research goal of interest, we could design different high-throughput calculations or experiments to generate the data we need, without being obsessed with existing data. Fortunately, many open-source software platforms facilitate the conduct of high-throughput calculations.[72–75] For high-throughput experiments, specific secondary development of experimental instruments is generally required, including sample preparation and performance testing. No matter how the data are obtained, it is usually necessary to assess and clean the data before presenting, by fixing or removing some incorrect, incomplete, duplicate, or corrupted data.[45,76] Due to the existence of abnormal values, the regularity mining of algorithms will be disturbed and thus lead to an unreliable prediction. Thus, data assessment and cleaning is an important prerequisite for the reliability of the ML model.[44,77,78]

However, due to the diversity of materials data, there has been no standardized way to handle this process. In addition to obtaining a large amount of data, data sampling and partition is also closely related to the reliability of the ML model. Flawed data sampling and partition are highly possible to cause bias. Some bias reduction methods will be discussed in Section 2.5. In short, ML, to some extent, has turned to the challenge of materials data.

2.3 | Featurization

Featurization or feature engineering is the step in which some features (or descriptors and fingerprints) are constructed from original data to describe different characteristics of materials, including the crystal/molecular composition and structure. The features in ML have strongly affected the model performance.[79–81] For example, establishing improved physics-based descriptors of perovskites enables a more accurate prediction of bandgap (E_b).[82] Thus, a critical challenge in applying ML to study materials problems is the selection of appropriate features.[83,84] In general, the construction of features is highly dependent on the goal of ML and involves a lot of knowledge and experience in the field of materials science.[85] For example, atomic radii and electronegativity are commonly used features in many ML works because they are key factors affecting dielectric/electrical properties based on some mature theories, such as the Goldschmidt tolerance factor and Pauling rules. This method of domain knowledge embedding makes the ML model more interpretable.
Meanwhile, existing human theories and knowledge inevitably introduce bias, which may result in nonoptimal results and even cover up new findings. In addition to artificially constructing the features, automatically extracting features from original data without human intervention may raise the possibility of discovering new underlying features beyond human cognition. This down-extraction process driven by data could be realized by some advanced algorithms, such as least absolute shrinkage and selection operator (LASSO) and genetic algorithms, and so forth.\cite{53,86} However, the drawback of the data-driven method is that the features may vary from different hyperparameters of extracting algorithms and not well explain the causality.\cite{87,88} To increase the efficiency and reliability of ML, feature selection and optimization are always performed where inessential or low-quality features would be removed or improved into new ones. For example, principal component analysis (PCA) is widely used to screen proper features.\cite{69} In short, ideal features are to maintain both the physical/chemical interpretability and the possibility of exploiting new effects.

\section{2.4 \hspace{1em} Algorithm}

The famous “no free lunch” theorem suggests that no algorithm is best for all possible learning situations and datasets. Selecting algorithms needs to be based on the goal of ML and available data or features. According to the learning style, ML algorithms could be categorized into three types: supervised learning, unsupervised learning, and reinforcement learning. In materials science, supervised learning (mainly including regression and classification) is the most common type to learn the mapping between input features (e.g., composition and structure) and output properties (e.g., bandgap, dielectric constant, and electrical conductivity). According to the core idea and functionality of algorithms, they could be further divided into more detailed concrete algorithms, such as linear regression, logistic regression, decision tree, K-nearest neighbor, neural network, support vector machine, random forest, and so forth. A detailed description of each algorithm could refer to several excellent textbooks and reviews.\cite{87,90} In a word, algorithm, as the executive of ML model, should be chosen after carefully considering multiple factors, such as goal, data, and features.

However, after selecting a suitable algorithm, it is not guaranteed that the model could exhibit good performance. The hyperparameters in algorithms define model configuration and architecture, which have an ultimate impact on the training results. For example, in neural networks, the number of hidden layers and the number of nodes in each layer are typical hyperparameters that let one control the training process. Thus, to achieve the best performance, it is important to understand how to optimize them. Here are some common strategies for tuning hyperparameters, such as grid search, random search, particle swarm optimization, Bayesian optimization, and gradient-based optimization. Thus, the selection and rational use of algorithms are also key processes in ML.

\section{2.5 \hspace{1em} Evaluation}

In practice, even if we obtain high-quality data and choose a suitable algorithm, it is inevitable that there are some prediction errors, including bias, variance, and noise, resulting in underfitting or overfitting. Underfitting refers to the model that cannot capture the underlying trend of the data, leading to high bias and low variance. While for overfitting, it generally means that the model corresponds too closely to the training data set, resulting in failing to accurately fit additional data set. Especially, when the ML model is too complex with many empirical parameters, it is prone to overfitting. To avoid the above issues, we could deal with them from several aspects, including feature engineering, model complexity, data set, and evaluation methods. Thus, during the training process of ML, we should evaluate the model iteratively to obtain better model performance. First, how to get a reasonable partition scheme of available data is the basis of algorithm reliability. The data are usually split into three: training data set, validation data set, and testing data set. Some effective evaluation methods are developed to reduce ML bias by partitioning data samples with high quality, such as the cross-validation method, the hold-out method, and the bootstrapping method.\cite{91} Then, when evaluating the accuracy of prediction results compared with training data, different evaluation indices are presented. Generally, for regression models, the means absolute error, mean square error, root mean square error, and the correlation coefficient ($R^2$) are common evaluation indices. For classification models, some indices, including the precision, the recall, the confusion matrix, and the accuracy are typically applied. If the evaluation of model performance is poor, we can constantly adjust the algorithms, features, and even data set until we get satisfactory results. Thus, evaluation could be regarded as the supervisor of ML as it helps in quantifying the model performance and guiding the improvement of the model.
2.6 | Application

Through continuous model evaluation and feedback optimization, trained ML models can be applied to materials research according to the research goal. For example, we could use ML to predict properties, discover new materials, optimize experimental design, and so on. However, good model performance is not equivalent to successfully solving scientific problems. How to further interpret and verify the ML results is the last critical step, which is an important link to combine theory with practice. In materials science, the crux of materials discovery and design is to identify the mutuality of four materials elements: composition/structure, property, synthesis/processing, and performance. Thus, excavating underlying relations and establishing new theories are vital problems for ML to address. Good interpretability of the ML model is helpful to the revelation of fundamental insights into the problem and the rationalization of design principles. On the other hand, on the basis of good model performance, how to experimentally verify the reliability of ML prediction results is a tricky business. Due to the complexity and variability of the experimental preparation process, many theoretically predicted materials are unable to be synthesized as expected, leading to a huge gap between theories and experiments. Thus, more and more attention is being paid to the integration of ML into the experimental process, such as the screening of materials composition, optimization of preparing technology, and evaluation of device performance. In summary, the applications of ML have brought great convenience to the R&D of energy storage materials but also face diverse challenges. More specific applications of ML will be discussed in the following section.

3 | ML-BASED MATERIALS RESEARCH AND DEVELOPMENT

In this section, we will review recent advances in the applications of ML on the R&D of energy storage materials, taking DCs and LIBs as examples. As summarized in many reviews, ML in materials science enables broad applications from the predictions of new materials/properties/theories to the optimization of composition/structure/process. Here, we mainly focus on three attractive aspects: discovering and designing novel materials, enriching theoretical simulations, and assisting experimentation and characterization.

3.1 | Discovering and designing novel materials

By combining different elements, compositions, and crystal/molecular structures, there is tremendous searching space for us to discover new materials with desired properties. However, the labor-intensive experimental process of trial and error is time-consuming, costly, and inefficient, which slows the pace of the R&D of energy storage materials. By benefitting from the improvement of computing techniques and algorithms, ML has shown great potential in accelerating the discovery of novel energy storage materials, such as dielectrics with high dielectric constant or high breakdown strength, solid electrolytes with high ionic conductivity, and so forth. Here, we present several representative research works to demonstrate how ML can predict properties and screen novel materials.

The dielectric constant $\varepsilon$ is a critical parameter in the design of polymer DCs. However, limited by the poor thermal stability (or low glass transition temperature $T_g$) of polymer dielectrics, finding a polymer dielectric with desired $\varepsilon$ and $T_g$ remains a challenge. To accelerate the discovery of satisfying polymer dielectrics, an ML-based model has been recently developed to instantly predict the frequency-dependent $\varepsilon$ and $T_g$ of polymers. As shown in Figure 3A, the training data set consisted of 1210 experimentally measured $\varepsilon$ at different frequencies and $T_g$. Then, by performing a fingerprinting scheme and the Gaussian process regression algorithm, the model was utilized to predict the $\varepsilon$ across the frequency range $60\text{–}10^{15}\text{Hz}$ and $T_g$ of synthesizable 11 000 candidate polymers, as presented in Figure 3B. Finally, taking the desired $\varepsilon$ and $T_g$ as screening criteria, five representative polymers with $\varepsilon > 5$ and $T_g > 450$ K were selected with the potential for high-temperature capacitors, as shown in Figure 3C. This study demonstrates that ML could be successfully used to rapidly predict properties and discover novel materials with desired property requirements.

As another important parameter determining the reliability and maximum operating voltage of DCs, the breakdown strength has gained lots of attention when designing dielectric materials. However, the complex dielectric breakdown mechanism is still not well understood, and thus how to discover and design novel dielectrics with high breakdown strength is confusing. To explore the intrinsic factors on breakdown strength and find novel dielectrics with high intrinsic breakdown strength, a transferable ML model was established to screen high-breakdown-strength dielectrics from 18928 ABX$_3$ perovskites, as shown in Figure 3D. First, based
on the dynamic stability of the perovskite crystal structure, 209 insulators were systematically selected. Then, based on the training data set from first-principle computations, the LASSO approach was used to find the explicit function of the intrinsic breakdown strength with the two most relevant features of bandgap and phonon cutoff frequency. Finally, the intrinsic breakdown strengths of all selected materials were predicted and three perovskite materials (SrBO₂F, SrBO₂F, and BSiO₂F) were identified as promising and worthy of further in-depth studies. In addition, ML could also play a guiding role in screening and designing polymer-based dielectric composites.[110–112] For example, as shown in Figure 4A, a powerful theoretical framework combined phase-field simulation and ML was established to study the electric–thermal–mechanical breakdown process and screen several suitable nanofillers for optimizing the breakdown strength.[113] First, high-throughput simulations were performed for the poly (vinylidene fluoride-hexafluoropropylene) (P(VDF-HFP))-based nanocomposites filled with nanoparticles of different properties. Then, taking the high-throughput phase-field simulation results as the database, ML was conducted to produce an analytical expression for predicting the breakdown strength by parameterizing the dielectric constant, electrical conductivity, and Young’s modulus. Based on screening results in Figure 4B, it was found that some nanocomposites filled with oxides such as Al₂O₃, SiO₂, MgO, and TiO₂ could exhibit higher breakdown strength, which was also verified by targeted experiments. This ML work develops a universal workflow of materials screening and design for optimizing the performances of dielectric nanocomposites.
For LIBs, ML is also contributing to accelerating the screening of novel battery materials,\textsuperscript{[114–117]} such as active electrodes and solid/liquid electrolyte materials as illustrated in Figure 1C. For example, in the case of solid electrolytes, massive efforts are focused on the search for inorganic/polymer solid ion conductors with high ionic conductivities and suitable mechanical properties.\textsuperscript{[118–122]} As shown in Figure 5A, an unsupervised learning approach was proposed based on unsupervised clustering of Li-containing compounds, and then 16 new fast Li conductors with room temperature conductivities of $10^{-4}$–$10^{-1} \text{S cm}^{-1}$ were predicted by ab initio molecular dynamics simulations.\textsuperscript{[123]} In this study, this unsupervised scheme was achieved by utilizing a limited quantity of conductivity data to prioritize a candidate list to increase the accuracy of screening results. In particular, three new materials systems, Li$_8$N$_2$Se, Li$_6$KBiO$_6$, and Li$_5$P$_2$N$_5$, exhibit room temperature conductivities exceeding $10^{-2} \text{S cm}^{-1}$, which are higher than that of the best known solid Li-ion conductors. Compared with inorganic solid electrolytes, polymer electrolytes possess the advantages of safety, cost, and flexibility, which have been regarded as one of the most promising electrolyte materials for LIBs.\textsuperscript{[125,126]} To explore the polymer electrolytes with higher ionic conductivity, a transfer-learned graph neural network was constructed to accurately predict the ionic conductivities of $10^4$ polymers in the polymer database, as illustrated in Figure 5B.\textsuperscript{[124]} It was found that the superionic conductors composed of charge-transfer complexes of aromatic polymers could exhibit the ionic conductivity of around $10^{-3} \text{S cm}^{-1}$ at room temperature. Contrary to the traditional concept of rubbery polymer electrolytes, this glassy design was proposed to achieve fast, decoupled motion of ionic species from polymer chains and thus enhance mechanical and thermal stability. This study is an excellent example to show the great potential of ML in discovering new materials without being fettered by human experience. In recent years, polymer blend or polymer-based composite electrolytes have been extensively studied to modulate the trade-off between ionic conductivity and mechanical property.\textsuperscript{[127–130]} Recently, an ML method has been employed to realize the multi-objective design of polymer blend electrolytes in two terms of ionic conductivity and viscosity,\textsuperscript{[131]} and it was concluded that polymer blend electrolytes with unequal molecular weights hardly improve electrolyte performance. Overall, plenty of research works suggests the potential of ML to discover new phenomena and novel materials, which tremendously promotes the breakthrough and innovation of energy storage materials.\textsuperscript{[97,105,132,133]}

In summary, plenty of successful works indicates that ML has been a powerful tool in the discovery and design of novel energy storage materials. At the same time, materials scientists are also strongly interested in figuring out the underlying mechanism behind novel
properties. As summarized in Section 2, by constructing some physical/chemical descriptors, ML results could become readable and understandable by summarizing some qualitative trends. However, it is still a huge challenge to accurately establish mathematical expressions. In the future, the ML model combined with physics and mathematics may promote the birth of new physical theories and then achieve innovative breakthroughs in new materials.

### 3.2 Enriching theoretical simulations

Over the last few decades, theoretical simulations have been viewed fundamentally as an endeavor to describe constitutive relationships in the fields of materials physics, mechanics, chemistry, and materials engineering. Up to now, there are lots of well-developed theories that could be used to study various materials phenomena at different scales of time and length, such as quantum mechanics, molecular dynamics, phase-field theory, and effective medium theory. However, existing computational simulations generally consume high computational costs and time for complex or multiscale problems. The implementation of ML could improve the efficiency and accuracy of theoretical simulations, and explore novel insights based on simulation results. In this section, we aim to illustrate how ML can enrich computational simulations by several representative examples from the atomic/molecular level to the mesoscopic/macroscopic level.

For the simulations at the atomic/molecular level, ML has been widely applied to improve the calculation accuracy or efficiency of density functional theory and molecular dynamics. For example, an efficient gradient-domain ML was developed to construct accurate energy-conserving molecular force fields of ab initio
molecular dynamics simulation.\textsuperscript{[148]} Moreover, Bayesian optimization was used to determine the optimal Hubbard $U$ parameters to improve the accuracy of density functional theory with a Hubbard $U$ correction.\textsuperscript{[149]} Except for the enhancement of algorithms and models, ML also can help to effectively discover more hidden information.\textsuperscript{[150,151]} For example, to motivate the development of novel highly conductive solid polymer electrolytes, molecular dynamics provide immense opportunities to study the role of molecular behaviors on ion transport properties. However, to fully understand what governs the ion mobility in solid polymer electrolytes and achieve global optimization is beyond the current capacity of conventional fully atomistic simulations. Thus, a new design approach that integrates coarse-grained molecular dynamics with ML was proposed to design highly conductive polymer electrolytes, as the framework shown in Figure 6.\textsuperscript{[152]} First, a continuous space with physical tunable and interpretable descriptors was constructed by the coarse-graining of chemical species. Then, the authors used Bayesian optimization to efficiently explore the relationships between the molecular descriptors and the conductivity of solid polymer electrolytes. It was concluded that changing TFSI$^-$, introducing secondary sites, and replacing PEO backbone chains are beneficial to improve the ionic conductivity of poly(ethylene oxide)-lithium bis(trifluoromethanesulfonyl)imide (PEO-LiTFSI) system. This method not only greatly reduces the computation cost and time but also gains useful mechanistic insights and optimize functions of solid polymer electrolytes.

At the mesoscopic level, the microstructure is one of the most important factors influencing the performance of energy storage materials,\textsuperscript{[108,153–156]} such as crystallization behavior and composite structure. Due to the complexity and variability of the microstructures in materials, comprehensively understanding the relationships between microstructures and properties is still a challenge. The combination of ML and mesoscale simulations is becoming an effective tool to tackle this issue. For example, to address microstructure–property relations of polymer nanocomposites for dielectrics energy storage, a materials development paradigm was developed by combining high throughput phase-field simulations and ML, as shown in Figure 7.\textsuperscript{[157]} Based on 6615 phase-field simulation results, an ML strategy was then performed to evaluate the capability of energy storage by a scoring function. The screening results revealed that taking parallel perovskite

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig6}
\caption{Framework of the coarse-grained molecular dynamics-Bayesian optimization (CGMD-BO). This workflow starts with the process of transforming the discrete conventional chemical species space to continuous space with coarse-graining (CG) parameters (\textsuperscript{①}–\textsuperscript{②}). Then, Bayesian optimization (BO) is employed to predict the relationships between the transport properties and the associated CG parameters (\textsuperscript{②}–\textsuperscript{③}). Reproduced with permission.\textsuperscript{[153]} Copyright 2020, American Chemical Society}
\end{figure}
nanosheets (e.g., Sr$_2$Ta$_3$O$_{10}$, Ca$_2$Nb$_3$O$_{10}$, LaNb$_2$O$_7$) as the nanofillers is beneficial to the improvement of the breakdown strength of polymer nanocomposites. Based on the guidance of ML results, a polymer nanocomposites P(VDF-HFP)/Ca$_2$Nb$_3$O$_{10}$ with an ultrahigh discharged energy density of 35.9 J cm$^{-3}$ at 853 kV m$^{-1}$ was fabricated. This study is a good combination of ML and phase-field models to realize the optimization and design of the composite structure.

Combined with theoretical simulations at the macroscopic level, ML is also applied to provide important guidelines on the safety assessment of LIBs for electric vehicles.$^{[158-160]}$ In the crash of electric vehicles, the battery pack may be damaged by the tremendous collision force, resulting in the electric short circuit, thermal runaway, and possible fire or explosion. However, the understanding of mechanical failure has been a blind spot for the battery community. Thus, it is an urgent need to identify the range of mechanical loading conditions ensuring the safe operation of battery cells, known as the “safety envelope.” To overcome the challenge of insufficient mechanical test data, a “bottom-up” modeling approach was proposed to include constitutive models, finite element models, and ML models, as shown in Figure 8.$^{[161]}$ First, a high-accuracy finite element model of a pouch cell was employed to generate ∼2500 simulations and then the ML model was used to predict the safety envelope. As shown in Figure 8D, on the one hand, it could quantitatively tell the maximal amount of deformation before failure by the constitutive models. On the other hand, this approach could quickly identify whether there is a short circuit at different mechanical loading conditions, as plotted in Figure 8F. The combination of finite element simulations and ML not only reduces the number of computations but also provides a high-accuracy prediction of macroscopic performance.

Therefore, ML could enrich various theoretical simulations at different isolated scales, including optimizing potential function or parameters, accelerating the algorithms, reducing the number of calculations, analyzing and interpreting the simulation results, and exploring new insights. However, many challenges remain in multiscale simulations of one system from microscale to macroscale. Given the powerful learning ability of ML, we believe that ML has the promising potential in connecting different theories to realize automatic multiscale simulations.$^{[135,162]}$
3.3 Assisting experimentation and characterization

The traditional experimental process highly depends on human experience and intuition, resulting in a slow and expensive cycle of the R&D of energy storage materials.[48] Furthermore, materials research has put forward higher and higher requirements for experimental characterization technology, which are gradually beyond the capabilities of existing advanced technologies. To overcome those limitations, researchers are trying to apply ML to assist experimentation and characterization, such as exploring optimal compositions, optimizing experimental processing, enhancing characterization technologies, and so forth.[92,163–169] Here, we will highlight several successful examples to spark more thinking on exploiting the potential of ML in experimentation and characterization.

In terms of experimental optimization design, it is attractive to employ ML to guide experimental synthesis to search for suitable materials with better properties.[92,93,170] Recently, to find better energy storage density of Pb-free BaTiO$_3$-based dielectrics at low electric fields as few experiments as possible, an ML-based approach was proposed to search optimal compounds.[171] As shown in Figure 9A, a closed feedback loop from experiments to ML was developed to guide the experimental design. It was started with physics intuition to reduce the search space in the (Ba$_{1-x-y}$Ca$_x$Sr$_y$)(Ti$_{1-u-v-w}$Zr$_u$Sn$_v$Hf$_w$)O$_3$ ceramics system. It is well known that the ceramics located around the crossover region in the composition–temperature phase diagram between relaxor ferroelectric phase and ferroelectric phase are more likely to exhibit better energy storage performance. Thus, the first step was to classify the compounds by predicting the phase diagram, as the example of elements Zr and Ca shown in Figure 9B. Then, taking the preselected compounds in the crossover region as the next search space, a regression model based on 182 experimental data was employed to make predictions. The prediction results of each ML would be verified by experiments, and then the experimental results would be fed back to the training data for iterative training. By performing only two active learning loops, the largest energy storage density $\approx 73$ mJ cm$^{-3}$ at 20 kV cm$^{-1}$ was found in the compound (Ba$_{0.86}$Ca$_{0.14}$)
(Ti$_{0.79}$Zr$_{0.11}$Hf$_{0.10}$)$_3$O$_3$, which is improved by 14% compared to the best in the training data, as shown in Figure 9C. This study provides an exemplary framework of ML to accelerate the searching process of multi-component materials in experiments.

Besides, based on partial experimental results, ML is also widely used to accurately predict the lifetime or safety of complex and nonlinear systems such as LIBs.

For example, the ML method was applied to assess the battery cycle life. First, the cycle lives of 124 commercial lithium iron phosphate/graphite cells were tested under fast-charging conditions to generate a training data set. Taking discharge voltage curves to exhibit capacity degradation, the ML model quantitatively predicted cycle life with 9.1% test error and classified cycle life into two groups with 4.9% test error. In addition, the degradation patterns of LIBs were identified using Gaussian process ML. Over 20,000 electrochemical impedance spectroscopy of commercial LIBs were measured at different states of health, charge, and temperatures. Without feature engineering, this model could accurately predict the remaining useful life. In addition, ML could be also used to guide the battery cell manufacturing process, which is difficult to plan and control by human experience.

Another particularly promising area of ML in assisting the R&D of energy storage materials is the enhancement of characterization techniques, including microscopy processing and spectroscopy analysis. Here, we mainly focus on recent applications of ML on image analysis to demonstrate the power of ML on recognition and classification. In materials science, it is highly desirable to characterize the local behaviors from nanoscale to microscale, which are closely correlated to the macroscopic properties of materials. By benefitting from the advancements of characterization techniques, such as scanning electron microscope, transmission electron microscope, atomic force microscope (AFM), and tomography, some local information from the atomic level to microstructure could be directly measured. However, plentiful characterization results are not well understood and utilized by human experience alone. ML could be an effective tool to assist in analyzing the characterization results, such as resolution enhancement, real-time processing,
identification and segmentation, reconstruction, and so forth.\textsuperscript{154,182–185} An interesting application of ML is constructing an artificial intelligence atomic force microscope (AI-AFM) system.\textsuperscript{186} As shown in Figure 10A, this system is capable of not only pattern recognition and feature identification in electrochemical systems and ferroelectric materials but can also classify via adaptive experimentation with additional probing at critical locations like domain wall and grain boundaries. The innovation in this study was using a support vector machine algorithm to realize high fidelity pixel-by-pixel recognition and classification in real time, which could be able to efficiently analyze AFM results without human interference. In addition, in LIBs, the microstructure evolution in a composite electrode highly determines the battery performance during the charging and discharging process. However, it is a frontier challenge to experimentally characterize the particle behavior in the conductive matrix and thus understand the electrochemical response under operating conditions. To address this issue, an ML-revealed statistics approach was proposed to study the behaviors of active particles, the carbon/binder domain, and pore structures in an Ni-rich LiNi\textsubscript{0.8}Mn\textsubscript{0.1}Co\textsubscript{0.1}O\textsubscript{2} composite cathode.\textsuperscript{187} To achieve the statistical representativeness of complex microstructures with more than 650 active particles, an ML model was developed to segment and label with superior accuracy and efficiency. As shown in Figure 10B, compared with the map results of a few representative particles obtained by phase-contrast tomography and traditional segmentation, the ML approach significantly improved the performance of identification. Up to now, extensive research has demonstrated the huge potential of ML in improving the accuracy or efficiency of experimental characterization.\textsuperscript{164,188–190}

In short, ML has already been widely integrated with the full life of experimental processes and shows promising potential for tackling complex problems. The application domains of ML in assisting experimentation and characterization range from materials preparation process and microstructure characterization to performance evaluation.

\textbf{FIGURE 10} (A) The concept of an artificial intelligence atomic force microscope (AI-AFM) with the functions of pattern recognition and classification without human interference. Reproduced with permission.\textsuperscript{186} Copyright 2018, RSC. (B) Comparison of the machine-learning-assisted segmentation results (B2) and traditional segmentation (watershed and separation algorithms) results (B3) for a few representative particles visualized by the hard X-ray phase-contrast nano-tomography technique (B1). Reproduced with permission.\textsuperscript{187} Copyright 2020, Springer Nature
Although ML has made good progress in facilitating experimental progress, there are still some challenges that need to be addressed, such as multiobjective optimization\cite{191,192} and inverse design.\cite{193–195}

4 | SUMMARY AND OUTLOOK

Mainly focusing on the energy storage materials in DCs and LIBs, we have presented a short review of the applications of ML on the R&D process. It should be pointed out that ML has also been widely used in the R&D of other energy storage materials, including fuel cells,\cite{196–198} thermoelectric materials,\cite{199,200} supercapacitors,\cite{201–203} and so on. So far, considerable progress has been made on the implementation of ML for discovering and designing novel materials, enriching theoretical simulations, and assisting experimentation and characterization, which has tremendously accelerated the R&D pace of energy storage materials. However, it is worth noting that numerous challenges still remain to be addressed. Here, we present our perspectives on some nascent but promising directions where ML may lead to further breakthroughs.

(1) Establishing full lifecycle materials database. As we reviewed above, most materials databases mainly consist of information about atom/molecule/crystal structure and corresponding intrinsic properties, which are inherent genes that fundamentally determine materials properties. However, materials performance could also be influenced by external factors during the full life cycle, such as the preparation process and service environment. To tackle this issue from a data-driven perspective, the urgent task is to build a full life-cycle materials database by collecting various types of data generated from the preparation and characterization process to the service process. Then, the key step is to establish a knowledge graph of composition–microstructure–preparation–service performance, which will be the basis for further mechanism analysis and optimization design. Among them, microstructure, as the extrinsic gene of materials, is very sensitive to the material composition, preparation process, service environment, and performance, and bridges all links in the full life-cycle process. However, insufficient attention is paid to studying the microstructural evolution via ML. We believe that taking the microstructure as the main line to series the full life-cycle research could not only provide more fertile ground for scientific innovation but also be more conducive to the practical use of materials.

(2) Developing multiobjective optimization algorithms. Whether a material can be used in engineering applications often requires comprehensive performance meeting the practical requirements. For example, for the application of a dielectric material, it should simultaneously meet the needs of multiple performance metrics instead of a single performance index, including the energy density, efficiency, working voltage, thermal stability, loss, and so forth. However, materials performances are affected by various complex factors, and there may exist an inverted coupling relationship among different performances, such as dielectric constant and breakdown strength for dielectric materials, ionic conductivity, and mechanical strength for solid electrolytes. Collaborative optimization of multiple performance indexes is a growing public concern in material research. Up to now, ML has made successful progress in optimizing single-objective performance as discussed in this review, such as screening solid electrolytes with high ionic conductivity or dielectrics with high dielectric constant. However, when dealing with two or more performances, the optimization problems will become tricky. Therefore, besides Pareto optimization or Bayesian optimization, new efficient algorithms urgently need to be developed.

(3) Improving quantifiable interpretability. Although many ML algorithms enable quick and accurate prediction, the black-box nature makes us trouble in understanding the decisions of ML systems. In materials science, quantifying physical correlation or causality and further establishing quantified expression between inputs and outputs is critical to understanding the underlying mechanism of novel phenomena and further extending to guide more materials research. Thus, developing ML models with both high predictive ability and high interpretability has been the hot pursuit of materials scientists.\cite{87,113,204–207} Recently, physics-informed ML integrates data and mathematical models, which may provide a new insight to discover hidden physics and tackle high-dimensional problems.\cite{60} Altogether, keeping the balance of quantifiable interpretability and intelligent forecasting of ML is a daunting challenge that requires deep integration and collaboration across multiple disciplines.

ACKNOWLEDGEMENTS

This study was supported by the Basic Science Center Program of NSFC (Grant No. 51788104), Major Research Plan of NSFC (Grant No. 92066103), NSF of China (Grant No. 52002300), Major Program of NSFC (Grant No. 51790490), and Young Elite Scientists Sponsorship Program by CAST (Frant No. 2019QNRC001). LQC acknowledges the
generous support from the Hamer Foundation through a Hamer Professorship at Penn State.

CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

REFERENCES
[1] Kittner N, Lill F, Kammen DM. Energy storage deployment and innovation for the clean energy transition. Nat Energy. 2017;2:1.
[2] Chu S, Majumdar A. Opportunities and challenges for a sustainable energy future. Nature. 2012;488:294-303.
[3] Chu S, Cui Y, Liu N. The path towards sustainable energy. Nat Mater. 2017;16:16-22.
[4] Gür TM. Review of electrical energy storage technologies, materials and systems: challenges and prospects for large-scale grid storage. Energy Environ Sci. 2018;11:2696-2767.
[5] Obama B. The irreversible momentum of clean energy. Science. 2017;355:126-129.
[6] Liu C, Li F, Ma LP, Cheng HM. Advanced materials for energy storage. Adv Mater. 2010;22:E28-E62.
[7] Schmidt O, Melchior S, Hawkes A, Staffell I. Projecting the future leveled cost of electricity storage technologies. Joule. 2019;3:81-91.
[8] Hall PJ, Bain EL. Energy-storage technologies and electricity generation. Energy Policy. 2008;36:4352-4355.
[9] Shao Y, El-Kady MF, Sun J, et al. Design and mechanisms of asymmetric supercapacitors. Chem Rev. 2018;118:9233-9280.
[10] Whittingham MS. Materials challenges facing electrical energy storage. MRS Bull. 2008;33:411-419.
[11] Simon P, Gogotsi Y. Perspectives for electrochemical capacitors and related devices. Nat Mater. 2020;19:1151-1163.
[12] Zhang G, Li Q, Allahyarov E, Li Y, Zhu L. Challenges and opportunities of polymer nanodielectrics for capacitive energy storage. ACS Appl Mater Interfaces. 2021;13:37939-37960.
[13] Thackeray MM, Wolverton C, Isaacs ED. Electrical energy storage for transportation—approaching the limits of, and going beyond, lithium-ion batteries. Energy Environ Sci. 2012;5:7854.
[14] Liu X-J, Zheng M-S, Chen G, Dang Z-M, Zha J-W. High-temperature polyimide dielectric materials for energy storage: theory, design, preparation and properties. Energy Environ Sci. 2022;15:56-82.
[15] Luo H, Zhou X, Ellingford C, et al. Interface design for high energy density polymer nanocomposites. Chem Soc Rev. 2019;48:4424-4465.
[16] Li Q, Yao F-Z, Liu Y, Zhang G, Wang H, Wang Q. High-temperature dielectric materials for electrical energy storage. Ann Rev Mater Res. 2018;48:219-243.
[17] Yang L, Kong X, Li F, et al. Perovskite lead-free dielectrics for energy storage applications. Prog Mater Sci. 2019;102:72-108.
[18] Shen Y, Zhang X, Li M, Lin Y, Nan C-W. Polymer nanocomposite dielectrics for electrical energy storage. Natl Sci Rev. 2017;4:23-25.
[19] Fan L-Z, He H, Nan C-W. Tailoring inorganic–polymer composites for the mass production of solid-state batteries. Nat Rev Mater. 2021;6:1003-1019.
[20] Liu W, Song MS, Kong B, Cui Y. Flexible and stretchable energy storage: recent advances and future perspectives. Adv Mater. 2017;29:1603436.
[21] Castelvecchi D, Stoye E. Chemistry Nobel honours world-changing batteries. Nature. 2019;574:308.
[22] Winter M, Barnett B, Xu K. Before Li ion batteries. Chem Rev. 2018;118:11433-11456.
[23] Nitta N, Wu F, Lee JT, Yushin G. Li-ion battery materials: present and future. Mater Today. 2015;18:252-264.
[24] Yue L, Ma J, Zhang J, et al. All solid-state polymer electrolytes for high-performance lithium ion batteries. Energy Storage Mater. 2016;5:139-164.
[25] Zheng Y, Yao Y, Ou J, Li M, Luo D. A review of composite solid-state electrolytes for lithium batteries: fundamentals, key materials and advanced structures. Chem Soc Rev. 2020; 49:8790-8839.
[26] Li M, Lu J, Chen Z, Amine K. 30 years of lithium-ion batteries. Adv Mater. 2018;30:1800561.
[27] Fan E, Li L, Wang Z, et al. Sustainable recycling technology for Li-ion batteries and beyond: challenges and future prospects. Chem Rev. 2020;120:7020-7063.
[28] Tabor DP, Roch LM, Saikin SK, et al. Accelerating the discovery of materials for clean energy in the era of smart automation. Nat Rev Mater. 2018;3:5.
[29] Liu Y, Guo B, Zou X, Li Y, Shi S. Machine learning assisted materials design and discovery for rechargeable batteries. Energy Storage Mater. 2020;31:434-450.
[30] Aykol M, Herring P, Anapolsky A. Machine learning for continuous innovation in battery technologies. Nat Rev Mater. 2020;5:725-727.
[31] Mistry A, Franco AA, Cooper SJ, Roberts SA, Viswanathan V. How machine learning will revolutionize electrochemical sciences. ACS Energy Lett. 2021;6:1422-1431.
[32] Brown KA, Brittman S, Maccaferri N, Jariwala D, Celano U. Machine learning in nanoscience: big data at small scales. Nano Lett. 2019;20:2-10.
[33] Gu GH, Noh J, Kim I, Jung Y. Machine learning for renewable energy materials. J Mater Chem A. 2019;7: 17096-17117.
[34] Pollice R, dos Passos Gomes G, Aldeghi M, et al. Data-driven strategies for accelerated materials design. Acc Chem Res. 2021;54:849-860.
[35] Lombardo T, Duquesnoy M, El-Bousidy H, et al. Artificial intelligence applied to battery research: hype or reality? Chem Rev. 2021. doi:10.1021/acs.chemrev.1c00108
[36] Chen C, Zuo Y, Ye W, Li X, Deng Z, Ong SP. A critical review of machine learning of energy materials. Adv Energy Mater. 2020;10:1903242.
[37] Suh C, Fare C, Warren JA, Pyzer-Knapp EO. Evolving the materials genome: how machine learning is fueling the next generation of materials discovery. Annu Rev Mater Res. 2020;50:1-25.
[38] Le TC, Winkler DA. Discovery and optimization of materials using evolutionary approaches. Chem Rev. 2016;116:6107-6132.
[39] Moosavi SM, Jablonka KM, Smit B. The role of machine learning in the understanding and design of materials. J Am Chem Soc. 2020;142:20273-20287.
[40] Sanchez-Lengeling B, Aspuru-Guzik A. Inverse molecular design using machine learning: generative models for matter engineering. Science. 2018;361:360-365.
[41] Saal JE, Oliynyk AO, Meredig B. Machine learning in materials discovery: confirmed predictions and their underlying approaches. *Annu Rev Mater Res*. 2020;50:49-69.

[42] Bedolla E, Padierna LC, Castañeda-Priego R. Machine learning for condensed matter physics. *J Phys Condens Matter*. 2020;33:053001.

[43] Mater AC, Coote ML. Deep learning in chemistry. *J Chem Inf Model*. 2019;59:2545-2559.

[44] Morgan D, Jacobs R. Opportunities and challenges for machine learning in materials science. *Annu Rev Mater Res*. 2020;50:71-103.

[45] Liu Y, Zhao T, Ju W, Shi S. Materials discovery and design using machine learning. *J Materiomics*. 2017;3:159-177.

[46] Vasudevan R, Pilania G, Balachandran PV. Machine learning for materials design and discovery. *J Appl Phys*. 2021;129(7):070401.

[47] Correa-Baena J-P, Hipolgaonkar K, van Duren J, et al. Accelerating materials development via automation, machine learning, and high-performance computing. *Joule*. 2018;2:1410-1420.

[48] Jain A, Ong SP, Hautier G, et al. Commentary: The Materials Project: a materials genome approach to accelerating materials innovation. *APL Mater*. 2013;1011002.

[49] Carleo G, Cirac I, Cranmer K, et al. Machine learning and the physical sciences. *Rev Mod Phys*. 2019;91:045002.

[50] Karniadakis GE, Kevekdisid IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. *Nat Rev Phys*. 2021;3:422-440.

[51] Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science*. 2015;349:255-260.

[52] Mohri M, Rostamizadeh A, Talwalkar A. *Foundations of Machine Learning*. MIT Press; 2018.

[53] Wang AT-T, Murdock RJ, Kauwe SK, et al. Machine learning for materials scientists: an introductory guide toward best practices. *Chem Mater*. 2020;32:4954-4965.

[54] Chen A, Zhang X, Zhou Z. Machine learning: accelerating materials development for energy storage and conversion. *InfoMat*. 2020;2:553-576.

[55] Artrith N, Butler KT, Coudert F-X, et al. Best practices in machine learning for chemistry. *Nat Chem*. 2021;13:505-508.

[56] Samudrala S, Rajan K, Ganapathy Subramanian B. *Informatics for Materials Science and Engineering*. Butterworth-Heinemann; 2013:97.

[57] Agrawal A, Choudhary A. Perspective: Materials informatics and big data: realization of the “fourth paradigm” of science in materials science. *APL Mater*. 2016;4:053208.

[58] Himanen L, Geurts A, Foster AS, Rinke P. Data-driven materials science: status, challenges, and perspectives. *Adv Sci*. 2019;6:1900808.

[59] Zhang Y, Ling C. Understanding the physical metallurgy of the CoCrFeMnNi high-entropy alloy: an atomistic simulation study. *npj Comput Mater*. 2018;4:1.

[60] Chibani S, Coudert F-X. Machine learning approaches for the prediction of materials properties. *APL Mater*. 2020;8:080701.

[61] Juan Y, Dai Y, Yang Y, Zhang J. Accelerating materials discovery using machine learning. *J Mater Sci Technol*. 2021;79:178-190.
negative breast cancer: a real-world study. Nat Rev Mater. 2020;54:1–7.

[81] Ghiringhelli LM, Vybiral J, Levchenko SV, Draxl C, Scheffler M. Big data of materials science: critical role of the descriptor. Phys Rev Lett. 2015;114:105503.

[82] Li C, Hao H, Xu B, et al. Improved physics-based structural descriptors of perovskite materials enable higher accuracy of machine learning. Comput Mater Sci. 2021;198:110714.

[83] He J, Li J, Liu C, et al. Machine learning identified materials descriptors for ferroelectricity. Acta Mater. 2021;209:116815.

[84] Himanen L, Jäger MOJ, Morooka EV, et al. DScribe: library descriptors for ferroelectricity. J Phys Chem C. 2021;125:397–402.

[85] Schmidt J, Marques MR, Botti S, Marques MA. Recent advances and applications of machine learning in solid-state materials science. npj Comput Mater. 2021;7:1.

[86] Dong G, Liu H. Feature Engineering for Machine Learning and Data Analytics. CRC Press; 2018.

[87] Tao Q, Xu P, Miller W. Machine learning for perovskite materials design and discovery. npj Comput Mater. 2020;6:1.

[88] Dong G, Liu H. Feature Engineering for Machine Learning and Data Analytics. CRC Press; 2018.

[89] Chen C, Raccuglia P, Elbert KC, Adler PD, et al. Machine-learning assisted materials discovery using failed experiments. Nature. 2016;533:73–76.

[90] Ren F, Ward L, Williams T, et al. Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments. Science Advances. 2018;4:eaq1566.

[91] Attia PM, Grover A, Jin N, et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. Nature. 2020;578:397–402.

[92] Turetsky A, Wessel J, Herrmann C, Thiede S. Battery production design using multi-output machine learning models. Energy Storage Mater. 2021;38:93–112.

[93] Gao T, Lu W. Machine learning toward advanced energy storage devices and systems. Science. 2020;24:101936.

[94] Chen X, Liu X, Shen X, Zhang Q. Applying machine learning to rechargeable batteries: from the microscale to the macroscale. Angewandte Chemie International Edition. 2021;60:24354–24366.

[95] Wu B, Han S, Shin KG, Lu W. Application of artificial neural networks in design of lithium-ion batteries. J Power Sources. 2018;395:128–136.

[96] Sendek AD, Cheon G, Pasta M, Reed EJ. Quantifying the search for solid Li-ion electrolyte materials by anion: a data-driven perspective. J Phys Chem C. 2020;124:8067–8079.

[97] Cheng X-B, Zhao C-Z, Yao Y-X, Liu H, Zhang Q. Crystal structure of the S1 subunit N-terminal domain from DcCoV UAE-HKU23 spike protein. Chem. 2019;5:74–82.

[98] Sendek AD, Yang Q, Cubuk ED, Duerloo K-AN, Cui Y, Reed EJ. Holistic computational structure screening of more than 12 000 candidates for solid lithium-ion conductor materials. Energy Environ Sci. 2017;10:306–320.

[99] Zha J-W, Zheng M-S, Fan B-H, Dang Z-M. Polymer-based dielectrics with high permittivity for electric energy storage: a review. Nano Energy. 2021;89:106438.

[100] Dong G, Liu H. Feature Engineering for Machine Learning and Data Analytics. CRC Press; 2018.

[101] Gao T, Yuan JK, Yao SH, Liao RJ. Flexible nanodielectric materials with high permittivity for power energy storage. Adv Mater. 2013;25:6334–6365.

[102] Li Q, Chen L, Gadinski MR, et al. Flexible high-temperature dielectric materials from polymer nanocomposites. Nature. 2015;523:576–579.

[103] Chan H, Cherukara M, Loeffler TD, Narayanan B, Sankaranarayanan SKRS. Machine learning enabled autonomous microstructural characterization in 3D samples. npj Comput Mater. 2020;6:1.

[104] Hao X. A review on the dielectric materials for high energy-storage application. J Adv Dielectr. 2013;3:1330001.

[105] Palneedhi H, Peddigari M, Hwang GT, Jeong DY, Ryu J. High-performance dielectric ceramic films for energy storage capacitors: progress and outlook. Adv Funct Mater. 2018;28:1803665.

[106] Hao X, Song Z, Hao H, et al. Homogeneous/Inhomogeneous—structured dielectrics and their energy-storage performances. Adv Mater. 2017;29:1601727.

[107] Kim C, Pilania G, Ramprasad R. Machine learning assisted predictions of intrinsic dielectric breakdown strength of ABX3 perovskites. J Phys Chem C. 2016;120:14575–14580.

[108] Zhu MX, Deng T, Dong L, Chen JM, Dang ZM. Integrated multifunctional properties of polypropylene composites by employing three-dimensional flower-like MgO with hierarchical surface morphology. IET Nanodielectr. 2021;4:27–37.

[109] Zhu M-X, Yu Q-C, Song H-G, Chen T-X, Chen J-M. Rational design of high-energy-density polymer composites by machine learning approach. ACS Appl Energy Mater. 2021;4:14449–14545.

[110] Feng D, Chen G, Yan X, et al. Molecular pyrolysis process and gas production characteristics of 3-element mixed insulation oil under thermal fault. High Voltage. 2021;6:693–701.

[111] Kunitski M, Eicke N, Huber P, et al. Double-slit photo-electron interference in strong-field ionization of the neon dimer. Nat Commun. 2019;10:1.

[112] Fujimura K, Seko A, Koyama Y, et al. Accelerated materials design of lithium superionic conductors based on first-principles calculations and machine learning algorithms. Adv Energy Mater. 2013;3:980–985.

[113] Gao X, Liu X, He R, et al. Designed high-performance lithium-ion battery electrodes using a novel hybrid model-data driven approach. Energy Storage Materials. 2021;36:435–458.

[114] Okubo M, Ko S, Dwibedi D, Yamada A. Designing positive electrodes with high energy density for lithium-ion batteries. J Mater Chem A. 2021;9:7407–7421.

[115] Roman D, Saxena S, Robu V, Pecht M, Flynn D. Machine learning pipeline for battery state-of-health estimation. Nat Mach Intell. 2021;3:447–456.

[116] Fergus JW. Ceramic and polymeric solid electrolytes for lithium-ion batteries. J Power Sources. 2010;195:4554–4569.
Zhang X, Wang S, Xue C, et al. Self‐fish J, Wagner GJ, Keten S. Mesoscopic and multiscale modelling in materials.

[131] Wheatle BK, Fuentes EF, Lynd NA, Ganesan V. Design of highly conductive polymer electrolytes by machine learning.

[132] Zhao Q, Stalin S, Zhao C, et al. Adolescent haze–related knowledge level study: a cross-sectional survey with sensitivity analysis.

[133] Liu Y, Zhou Q, Cui G. Machine learning boosting the development of advanced lithium batteries.

[134] de Pablo JJ, Jackson NE, Webb MA, et al. New frontiers for the materials genome initiative.

[135] Fish J, Wagner GJ, Keten S. Mesoscopic and multiscale modelling in materials.

[136] Meng YS, Arroyo-de Dompablo ME. First principles computational materials design for energy storage materials in lithium ion batteries.

[137] Nolan AM, Zhu Y, He X, Bai Q, Mo Y. Computation‐accelerated design of materials and interfaces for all‐solid‐state lithium‐ion batteries.

[138] Chen L-Q. Phase‐field models for microstructure evolution.

[139] Chen L-Q, Zhao Y. From classical thermodynamics to phase-field method.

[140] Nan C-W. Physics of inhomogeneous inorganic materials.

[141] Louie SG, Chan Y-H, Felipe H, Li Z, Qiu DY. Discovering and understanding materials through computation.

[142] Friederich P, Häse F, Froppe J, Aspuru-Guzik A. Machine‐learned potentials for next-generation material simulations.

[143] Dunjko V, Briegel HJ. Machine learning & artificial intelligence in the quantum domain: a review of recent progress.

[144] Rao KK, Yao Y, Grabow LC. Accelerated modeling of lithium diffusion in solid state electrolytes using artificial neural networks.

[145] Wang C, Aoyagi K, Wisesa P, Mueller T. The complete chloroplast genome of Polygonatum odoratum (Liliaceae), an endemic medicinal herb.

[146] Guo H, Wang Q, Stuke A, Urban A, Artrith N. Evaluating the health risks of pneumonia from airborne bacterial communities using 16S rDNA sequences of pneumonia-related pathogens.

[147] Jackson NE, Bowen AS, Antony LW, Webb MA, Vishwanath V, de Pablo JJ. Electronic structure at coarse-grained resolutions from supervised machine learning.

[148] Chmiela S, Tkatchenko A, Sauceda HE, Poltavsky I, Schütt KT, Müller KR. Machine learning of accurate energy-conserving molecular force fields.

[149] Yu M, Yang S, Wu C, Marom N. Machine learning the Hubbard U parameter in DFT+U using Bayesian optimization.

[150] Dijkstra M, Luijten E. From predictive modelling to guided understanding microstructure evolution.

[151] Han C, Cherukara M, Loeffler TD, Narayanan B, Sankaranarayanan SKRS. Machine learning enabled autonomous microstructural characterization in 3D samples.

[152] Priya P, Aluru N. Accelerated design and discovery of perovskites with high conductivity for energy applications through machine learning.

[153] Liu Y, Zhou Q, Cui G. Machine learning boosting the development of advanced lithium batteries.

[154] de Pablo JJ, Jackson NE, Webb MA, et al. New frontiers for the materials genome initiative.

[155] Pan H, Kursunovic A, Lin Y-H, Nan C-W, Mac-Driscoll J.L. Dielectric films for high performance capacitive energy storage: multiscale engineering.
[156] Oh JM, Venters CC, Di C, et al. U1 snRNPs regulate cancer cell migration and invasion in vitro. Nat Commun. 2020;11:1.

[157] Shen Z-H, Bao Z-W, Cheng X-X, et al. Designing polymer nanocomposites with high energy density using machine learning. npj Comput Mater. 2021;7:1.

[158] Jia Y, Li J, Yuan C, et al. Data-driven safety risk prediction of lithium-ion battery. Adv Energy Mater. 2021;11:2003868.

[159] Finegan DP, Cooper SJ. Battery safety: data-driven prediction of failure. Joule. 2019;3:2599-2601.

[160] Ng M-F, Zhao J, Yan Q, Conduit GJ, Seh ZW. Predicting the state of charge and health of batteries using data-driven machine learning. Nat Mach Intell. 2020;2:161-170.

[161] Li W, Zhu J, Xia Y, Gorji MB, Wierzbicki T. Data-driven safety envelope of lithium-ion batteries for electric vehicles. Joule. 2019;3:2703-2715.

[162] Peng G, Alber M, Tepole AB, et al. Multiscale modeling meets machine learning: what can we learn? Archiv Comput Methods Eng. 2021;28:1017-1037.

[163] Green ML, Choi CL, Hattrick-Simpers JR, et al. Fulfilling the promise of the materials genome initiative with high-throughput experimental methodologies. Appl Phys Rev. 2017;4:011105.

[164] Chan H, Cherukara M, Loeffler TD, Narayanan B, Sankaranarayanan SK. Machine learning enabled autonomous microstructural characterization in 3D samples. npj Comput Mater. 2020;6:1.

[165] Chowdhury A, Kautz E, Yener B, Lewis D. Image driven machine learning methods for microstructure recognition. Comput Mater Sci. 2016;123:176-187.

[166] Bostanabad R, Zhang Y, Li X, et al. Computational microstructure characterization and reconstruction: review of the state-of-the-art techniques. Prog Mater Sci. 2018;95:1-41.

[167] Benayad A, Diddens D, Heuer A, et al. High-throughput experimentation and computational freeway lanes for accelerated battery electrolyte and interface development research. Adv Energy Mater. 2021;202678.

[168] Lyu Y, Liu Y, Cheng T, Guo B. High-throughput characterization methods for lithium batteries. J Materiomics. 2017;3:221-229.

[169] Granda JM, Donina L, Dragone V, Long D-L, Cronin L. Controlling an organic synthesis robot with machine learning to search for new reactivity. Nature. 2018;559:377-381.

[170] Balachandran PV, Kowalski B, Sehirlioglu A, Lookman T. Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. Nat Commun. 2018;9:1.

[171] Yuan R, Tian Y, Xue D, et al. Accelerated search for BaTiO3-based ceramics with large energy storage at low fields using machine learning and experimental design. Adv Sci. 2019;6:1901395.

[172] Chemali E, Kollmeyer PJ, Preindl M, Ahmed R, Emadi A. Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries. IEEE Trans Ind Electron. 2017;65:6730-6739.

[173] Finegan DP, Zhu J, Feng X, et al. The application of data-driven methods and physics-based learning for improving battery safety. Joule. 2021;5:316-329.

[174] Severson KA, Attia PM, Jin N, et al. Data-driven prediction of battery cycle life before capacity degradation. Nat Energy. 2019;4:383-391.

[175] Oh JM, Venters CC, Di C, et al. U1 snRNPs regulate cancer cell migration and invasion in vitro. Nat Commun. 2020;11:1.

[176] Nomura Y, Yamamoto K, Fuji M, Hirayama T, Igaki E, Saitoh K. Dynamic imaging of lithium in solid-state batteries by operando electron energy-loss spectroscopy with sparse coding. Nat Commun. 2020;11:1.

[177] Xie T, France-Lanord A, Wang Y, Shao-Horn Y, Grossman JC. Double-slit photoelectron interference in strong-field ionization of the neon dimer. Nat Commun. 2019;10:1.

[178] Zhang F, Williams KN, Edwards D, et al. Maximizing Information: a machine learning approach for analysis of complex nanoscale electromechanical behavior in defect-rich PZT films. Small Methods. 2021;2100552.

[179] Rabolovsky NM, Lednev IK. Towards development of a novel universal medical diagnostic method: Raman spectroscopy and machine learning. Chem Soc Rev. 2020;49:7428-7453.

[180] Meza Ramirez CA, Greenop M, Ashton L, Rehman I U. Applications of machine learning in spectroscopy. Appl Spectrosc Rev. 2021;56:733-763.

[181] Ellis DJ, Broadhurst D, Kell DB, Rowland JJ, Goodacre R. Rapid and quantitative detection of the microbial spoilage of meat by Fourier transform infrared spectroscopy and machine learning. Appl Environ Microbiol. 2002;68:2822-2828.

[182] Waller L, Tian L. Computational imaging: machine learning for 3D microscopy. Nature. 2015;523:416-417.

[183] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521:436-444.

[184] de Haan K, Ballard ZS, Rivenson Y, Wu Y, Ozcan A. Resolution enhancement in scanning electron microscopy using deep learning. Sci Rep. 2019;9:1.

[185] Akers S, Kautz E, Trevino-Gavito A, et al. Rapid and flexible segmentation of electron microscopy data using few-shot machine learning. npj Comput Mater. 2021;7:1.

[186] Huang B, Li Z, Li J. An artificial intelligence atomic force microscope enabled by machine learning. Nanoscale. 2018; 10:21320-21326.

[187] Oh JM, Venters CC, Di C, et al. U1 snRNPs regulate cancer cell migration and invasion in vitro. Nat Commun. 2020;11:1.

[188] Nelson CT, Ghosh A, Oxley M, et al. Deep learning ferroelectric polarization distributions from STEM data via with and without atom finding. npj Comput Mater. 2021;7:1.

[189] Hong QJ, Schroers J, Hofmann D, Curtarolo S, Asta M, van de Walle A. Theoretical prediction of high melting temperature for a Mo–Ru–Ta–W HCP multiprincipal element alloy. npj Comput Mater. 2021;7:1.

[190] Kaufmann K, Zhu C, Rosengarten AS, et al. Crystal symmetry determination in electron diffraction using machine learning. Science. 2020;367:564-568.

[191] Ashby M. Multi-objective optimization in material design and selection. Acta Mater. 2000;48:359-369.

[192] Gopakumar AM, Balachandran PV, Xue D, Gubernatis JE, Lookman T. Multi-objective optimization for materials discovery via adaptive design. Sci Rep. 2018;8:1.
Zunger A. Inverse design in search of materials with target functionalities. Nat Rev Chem. 2018;2:1.

Noh J, Kim J, Stein HS, et al. Inverse design of solid-state materials via a continuous representation. Matter. 2019;1:1370-1384.

Chen CT, Gu GX. Generative deep neural networks for inverse materials design using backpropagation and active learning. Adv Sci. 2020;7:1902607.

Ding R, Ding Y, Zhang H, et al. Applying machine learning to boost the development of high-performance membrane electrode assembly for proton exchange membrane fuel cells. J Mater Chem A. 2021;9:6841-6850.

Zou X, Pan J, Sun Z, et al. Machine learning analysis and prediction models of alkaline anion exchange membranes for fuel cells. Energy Environ Sci. 2021;14:3965-3975.

Zhang Z, Li S, Xiao Y, Yang Y. Intelligent simultaneous fault diagnosis for solid oxide fuel cell system based on deep learning. Appl Energy. 2019;233:930-942.

Wang T, Zhang C, Snoussi H, Zhang G. Machine learning approaches for thermoelectric materials research. Adv Funct Mater. 2020;30:1906041.

Recatala-Gomez J, Suwardi A, Nandhakumar I, Abutaha A, Hippalgaonkar K. Toward accelerated thermoelectric materials and process discovery. ACS Appl Energy Mater. 2020;3:2240-2257.

Ghosh S, Rao GR, Thomas T. Machine learning-based prediction of supercapacitor performance for a novel electrode material: cerium oxynitride. Energy Storage Mater. 2021;40:426-438.

Zhou Y, Huang Y, Pang J, Wang K. Remaining useful life prediction for supercapacitor based on long short-term memory neural network. J Power Sources. 2019;440:227149.

Zhou M, Vassallo A, Wu J. Data-driven approach to understanding the in-operando performance of heteroatom-doped carbon electrodes. ACS Appl Energy Mater. 2020;3:5993-6000.

Iwasaki Y, Sawada R, Stanev V, et al. Identification of advanced spin-driven thermoelectric materials via interpretable machine learning. npj Comput Mater. 2019;5:1.

Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019;1:206-215.

Dai M, Demirel MF, Liang Y, Hu J-M. Graph neural networks for an accurate and interpretable prediction of the properties of polycrystalline materials. npj Comput Mater. 2021;7:1.

Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B. Definitions, methods, and applications in interpretable machine learning. Proc Natl Acad Sci. 2019;116:22071-22080.

How to cite this article: Shen Z-H, Liu H-X, Shen Y, Hu J-M, Chen L-Q, Nan C-W. Machine learning in energy storage materials. Interdiscip Mater. 2022;1:175-195. doi:10.1002/idm2.12020