Journal of Multiscale Material Informatics
ISSN 3047-5724

Review Article

Machine Learning-Assisted Discovery and Optimization of Sodium-Ion Batteries: A Review

Gustina Alfa Trisnapradika[1], Harun Al Azies[2], Muhamad Akrom[3*], Usman Sudibyo[4], Noor Ageng Setiyanto[5]

[1,2,3,4,5] Research Group for Quantum Computing and Materials Informatics, Universitas Dian Nuswantoro, Semarang 50131, Indonesia
*Corresponding: m.akrom@dsn.dinus.ac.id

Abstract
Sodium-ion batteries (SIBs) have emerged as a promising alternative to lithium-ion batteries due to the natural abundance, low cost, and wide geographic availability of sodium resources. However, their practical implementation is hindered by challenges such as lower energy density, slower ion diffusion, and limited cycle stability. In recent years, machine learning (ML) has been increasingly applied to accelerate the discovery, design, and optimization of SIB materials and systems. This review provides a comprehensive overview of ML applications in sodium-ion battery research, including electrode material discovery, electrolyte optimization, performance prediction, and degradation analysis. Various ML techniques, such as supervised learning, unsupervised learning, and deep learning, are discussed in relation to their roles in materials informatics. Additionally, challenges such as data scarcity, model interpretability, and transferability are critically analyzed. Finally, future perspectives on integrating ML with high-throughput experiments and quantum computing are highlighted to guide next-generation sodium-ion battery research.

Keywords: Sodium-ion batteries; Machine learning; Electrode materials; High-throughput screening; Energy storage systems.

Received: 6 April 2026 / **Revised:** 29 April 2026 / **Accepted:** 4 Mei 2026 / **Published:** 5 Mei 2026

1. INTRODUCTION

The rapid transition toward renewable energy systems has intensified the demand for efficient, scalable, and cost-effective energy storage technologies. While lithium-ion batteries (LIBs) have dominated the energy storage market for decades due to their high energy density and mature technology, concerns about limited lithium resources, uneven global distribution, and rising costs have driven the search for alternative battery chemistries. In this context, sodium-ion batteries (SIBs) have emerged as a promising candidate, particularly for large-scale applications such as grid energy storage, where cost and resource availability are more critical than gravimetric energy density.

Sodium offers several intrinsic advantages over lithium, including its natural abundance, low extraction cost, and widespread availability across the Earth's crust. These characteristics make SIBs highly attractive for sustainable and economically viable energy storage solutions. Moreover, the electrochemical working principles of sodium-ion batteries are analogous to those of LIBs, allowing the adaptation of existing battery manufacturing infrastructure. Despite these advantages, the practical deployment of SIBs remains constrained by several fundamental challenges. The larger ionic radius of Na^+ compared to Li^+ leads to sluggish diffusion kinetics and significant structural strain in electrode materials during cycling. This often results in lower energy density, poor rate capability, and limited cycle life. Additionally, identifying suitable electrode and electrolyte materials that can reversibly host sodium ions with high stability remains a major bottleneck.

To address these challenges, conventional experimental approaches based on trial-and-error and incremental material optimization are increasingly insufficient given the vast chemical and structural search space. In recent years, machine learning (ML) has emerged as a transformative paradigm in materials science, enabling data-driven discovery and optimization of complex systems. By leveraging large datasets derived from experiments and first-principles calculations, ML models can rapidly predict key material properties, uncover hidden correlations, and guide the design of high-performance battery components. In the context of sodium-ion batteries, ML has been applied to various aspects, including electrode material screening, electrolyte design, performance prediction, and degradation analysis.

Despite the growing body of literature on ML-assisted battery research, most studies have primarily focused on lithium-ion systems, resulting in a significant knowledge gap in sodium-ion battery applications.

The direct transfer of ML models from LIBs to SIBs is often ineffective due to fundamental differences in electrochemical behavior, ionic size effects, and structural dynamics. Furthermore, challenges such as limited high-quality datasets, a lack of standardized data representation, and poor model interpretability hinder the broader adoption of ML in SIB research. These issues highlight the need for a systematic, critical review that specifically examines the role of machine learning in sodium-ion battery development.

This review aims to provide a comprehensive and critical overview of recent advances in applying machine learning to sodium-ion batteries. Unlike previous reviews that focus broadly on energy storage or on lithium-based systems, this work specifically emphasizes the challenges and opportunities related to SIBs. The review systematically discusses the fundamental principles of SIBs, various ML techniques employed in battery research, and their applications in materials discovery, performance optimization, and lifetime prediction. In addition, key challenges, including data scarcity, model generalization, and interpretability, are critically analyzed. Finally, future research directions, including the integration of ML with high-throughput experimentation and emerging paradigms such as quantum machine learning, are proposed to accelerate the development of next-generation sodium-ion battery technologies.

2. FUNDAMENTAL CONCEPTS

2.1 Working Principle

Sodium-ion batteries (SIBs) operate based on the reversible electrochemical intercalation and deintercalation of sodium ions (Na^+) between the cathode and anode through an electrolyte medium, following a mechanism that is fundamentally analogous to that of lithium-ion batteries. During the charging process, sodium ions are extracted from the cathode host structure and migrate through the electrolyte toward the anode, where they are stored via intercalation, adsorption, or alloying mechanisms, depending on the anode material. Conversely, during discharge, the stored sodium ions return to the cathode, generating an electric current through the external circuit. Although this mechanism appears similar to that of lithium-based systems, the larger ionic radius of Na^+ (≈1.02 Å) compared to Li^+ (≈0.76 Å) introduces significant differences in ion transport behavior, diffusion kinetics, and structural stability of host materials.

The performance of SIBs is highly dependent on the physicochemical properties of their constituent components, particularly the electrode materials and electrolyte system. Cathode materials typically include layered transition metal oxides, polyanionic compounds such as phosphates and sulfates, and Prussian blue analogs, each offering distinct advantages in terms of structural stability, operating voltage, and sodium diffusion pathways. However, these materials often suffer from phase transitions, structural distortions, and limited reversible capacity during long-term cycling. On the anode side, hard carbon has emerged as one of the most promising candidates due to its ability to accommodate sodium ions through a combination of adsorption in nanopores and intercalation between disordered graphene layers. Nevertheless, the sodium storage mechanism in hard carbon remains not fully understood, particularly regarding the formation of low-potential plateaus and irreversible capacity loss during the initial cycles.

The electrolyte plays a critical role in ensuring efficient ion transport and interfacial stability at the electrode-electrolyte interface. Conventional electrolytes for SIBs typically consist of sodium salts such as $NaPF_6$ or $NaClO_4$ dissolved in organic carbonate solvents. However, challenges such as electrolyte decomposition, unstable solid electrolyte interphase (SEI) formation, and safety concerns remain significant obstacles. The stability and composition of the SEI layer are particularly crucial, as they directly influence the battery's cycle life and coulombic efficiency. Unlike lithium-ion systems, the SEI in sodium-ion batteries tends to be less stable due to differences in reduction potentials and solvent interactions, further complicating performance optimization.

From a thermodynamic and kinetic perspective, the electrochemical performance of SIBs is governed by a complex interplay -among ion diffusion, charge transfer resistance, and the structural evolution of electrode materials. The slower diffusion kinetics of sodium ions often result in reduced rate capability, while repeated insertion and extraction processes can induce mechanical stress and structural degradation. These factors highlight the importance of designing materials with open frameworks, stable crystal structures, and favorable diffusion pathways. In this context, understanding the structure–property relationships at the atomic and mesoscale levels is essential for improving SIB performance, which also underscores the relevance of data-driven approaches such as machine learning in capturing these complex interactions.

2.2 Challenges in SIB Development

Despite their potential as a cost-effective alternative to lithium-ion batteries, sodium-ion batteries face several intrinsic challenges that limit their practical application. One of the most fundamental issues arises from the larger ionic radius and higher atomic mass of sodium, which lead to slower diffusion kinetics and lower specific capacity compared to lithium-based systems. This limitation is particularly evident in electrode materials, where the insertion of sodium ions often induces significant lattice strain and volume

expansion. Over repeated charge-discharge cycles, these structural changes can result in mechanical degradation, particle cracking, and eventual loss of electrochemical activity.

Another critical challenge is the relatively low operating voltage of sodium-ion batteries, which directly impacts their energy density. Since sodium has a higher redox potential than lithium, the overall cell voltage of SIBs is inherently lower, making it difficult to achieve energy densities comparable to those of LIBs. This limitation restricts the applicability of SIBs in high-energy-demand applications such as electric vehicles, although they remain promising for stationary energy storage systems. Furthermore, identifying suitable cathode materials that can deliver both high voltage and long-term stability remains an ongoing research challenge. Many high-capacity materials undergo irreversible phase transitions or structural instability during cycling, leading to rapid capacity fading.

In addition to electrode-related challenges, the development of stable, efficient electrolytes is a major concern. The formation of a robust, stable solid electrolyte interphase is more difficult in sodium-ion systems due to differences in electrochemical reactivity relative to lithium. This often results in continuous electrolyte decomposition, low coulombic efficiency, and poor cycle life. Moreover, safety issues related to flammable organic electrolytes further complicate the design of reliable SIB systems.

From a broader perspective, one of the most significant bottlenecks in advancing sodium-ion battery technology is the vast and complex design space of materials and system configurations. Traditional experimental approaches are often time-consuming and resource-intensive, making it impractical to explore all possible combinations of materials and conditions. Additionally, the lack of standardized and high-quality datasets further limits the ability to systematically optimize SIB performance. These challenges highlight the urgent need for advanced computational and data-driven methodologies, particularly machine learning, to accelerate the discovery of novel materials and provide deeper insights into the underlying mechanisms governing sodium-ion battery behavior.

3. MACHINE LEARNING TECHNIQUES

3.1 Supervised Learning

Supervised learning has become one of the most dominant machine learning paradigms in sodium-ion battery research due to its capability to establish quantitative relationships between material descriptors and electrochemical properties. In the context of SIBs, supervised models are typically trained on datasets derived from experimental measurements or density functional theory calculations, with input features including compositional descriptors, crystal-structure parameters, electronic properties, and thermodynamic stability indicators. These models are then used to predict key performance metrics, including specific capacity, average voltage, diffusion barriers, and cycling stability.

Unlike lithium-ion systems, where large and relatively standardized datasets are available, supervised learning in SIB research often operates under conditions of limited and heterogeneous data. This constraint necessitates careful feature engineering and model selection to avoid overfitting and ensure generalizability. Algorithms such as random forest and gradient boosting have been widely adopted due to their robustness against small datasets and their ability to capture nonlinear relationships without extensive hyperparameter tuning. Support vector machines have also shown effectiveness in classification tasks, such as distinguishing stable versus unstable electrode materials. Importantly, recent studies have begun incorporating domain knowledge into feature construction, such as ionic radius mismatch, lattice volume change, and redox-active species, which significantly improve predictive performance and physical interpretability. Nevertheless, the reliability of supervised learning models in SIB applications remains strongly dependent on data quality and diversity, highlighting a critical gap in curated sodium-specific datasets.

3.2 Unsupervised Learning

Unsupervised learning plays a crucial role in sodium-ion battery research by enabling the exploration of complex, high-dimensional materials data without labeled outputs. This is particularly important in SIB systems, where experimental data are scarce and often incomplete. Unsupervised techniques are commonly employed to identify hidden patterns, cluster materials with similar electrochemical behavior, and reduce the dimensionality of descriptor space to facilitate visualization and interpretation.

In SIB research, clustering algorithms have been used to categorize cathode and anode materials based on structural and compositional similarities, which helps in identifying families of materials with promising electrochemical characteristics. Dimensionality reduction methods such as principal component analysis and t-distributed stochastic neighbor embedding are frequently used to map high-dimensional feature spaces into lower-dimensional representations, enabling researchers to identify trends in ion diffusion pathways, structural stability, and voltage profiles. These techniques are particularly valuable in the early stages of materials discovery, where the goal is to narrow down a large candidate space into a manageable subset for further analysis.

Moreover, unsupervised learning can reveal underlying correlations that are not immediately apparent through conventional analysis, such as the relationship between structural disorder and sodium storage mechanisms in hard carbon. However, a key limitation of these methods is the difficulty in translating discovered patterns into actionable design rules, especially when physical interpretability is not explicitly incorporated. This highlights the need for hybrid approaches that combine unsupervised learning with domain-specific knowledge to enhance their practical relevance in SIB development.

3.3 Deep Learning

Deep learning has emerged as a powerful tool for modeling complex structure–property relationships in sodium-ion battery systems, particularly for unstructured or highly nonlinear data. In contrast to traditional machine learning methods that rely heavily on handcrafted features, deep learning models can automatically learn hierarchical representations from raw input data, making them especially suitable for capturing intricate atomic and electronic interactions.

In recent years, graph neural networks have gained significant attention in SIB research due to their ability to represent materials as graphs, where atoms are treated as nodes and interatomic bonds as edges. This representation allows the model to learn directly from crystal structures without explicit feature engineering. Graph-based models have demonstrated strong performance in predicting properties such as formation energy, voltage profiles, and ion diffusion barriers, which are critical for evaluating electrode materials. Their ability to generalize across different material classes makes them particularly attractive for exploring the diverse chemical space of sodium-ion systems.

Convolutional neural networks and other deep architectures have also been applied to analyze imaging data, such as electron microscopy images, to study morphological changes and degradation mechanisms in electrodes. Additionally, recurrent neural networks and transformer-based models are being explored for time-series analysis of battery cycling data, enabling more accurate prediction of battery lifetime and performance degradation.

Despite these advancements, the application of deep learning in sodium-ion batteries is still in its early stages compared to lithium-ion systems. One of the main challenges is the lack of sufficiently large, high-quality datasets needed to train deep models effectively. Furthermore, the black-box nature of deep learning limits interpretability, a critical requirement in scientific research, where understanding the underlying physical mechanisms is as important as prediction accuracy. Consequently, there is a growing interest in developing explainable deep learning models and integrating them with physics-based constraints to improve both reliability and insight.

4. ML APPLICATIONS IN SIB

4.1 Electrode Material Discovery

The discovery of high-performance electrode materials remains one of the most critical challenges in sodium-ion battery development, primarily due to the vast chemical space and the complex interplay between structural stability, ion diffusion, and electrochemical performance. In this context, machine learning has emerged as a powerful tool to accelerate the identification of promising cathode and anode materials by enabling rapid screening of large materials databases. Unlike traditional approaches that rely heavily on density functional theory calculations or experimental trial-and-error, ML-driven frameworks can predict key properties such as specific capacity, average voltage, formation energy, and diffusion barriers with significantly reduced computational cost.

Recent studies have demonstrated the effectiveness of combining high-throughput computational datasets with supervised learning models to identify novel sodium host materials. In particular, tree-based ensemble methods and kernel-based models have shown strong performance in predicting electrochemical properties from compositional and structural descriptors. More recently, graph neural networks have gained prominence due to their ability to directly learn from crystal structures, capturing local atomic environments and long-range interactions more effectively than conventional descriptor-based models. This is particularly relevant for sodium-ion systems, where subtle structural variations can significantly influence ion mobility and stability.

However, despite these advancements, a critical limitation persists in the generalizability of ML models across different material families. Many models are trained on relatively narrow datasets, often focusing on specific classes such as layered oxides or polyanionic compounds, which restricts their applicability to unexplored materials. Furthermore, the lack of standardized benchmarks for evaluating model performance in SIB-specific contexts makes it difficult to compare different approaches. These issues highlight the need for more diverse datasets and robust validation frameworks to ensure reliable materials discovery.

4.2 Electrolyte Optimization

The development of efficient and stable electrolytes is essential for improving the overall performance and safety of sodium-ion batteries, yet it remains a complex and underexplored area compared to electrode materials. Machine learning has recently begun to play an important role in this domain by enabling predictions of key electrolyte properties, including ionic conductivity, electrochemical stability window, viscosity, and compatibility at electrode interfaces.

Data-driven approaches have been employed to analyze large chemical spaces of solvent–salt combinations, allowing for the identification of optimal electrolyte formulations that balance conductivity and stability. In particular, supervised learning models trained on experimentally measured or simulated datasets have been used to predict ionic transport properties and guide the design of novel electrolyte systems. Additionally, ML has been applied to study the formation and stability of the solid electrolyte interphase, which is a critical factor influencing battery lifespan and coulombic efficiency.

Nevertheless, the application of ML in electrolyte optimization for SIBs is still relatively limited compared to its use in electrode discovery. One of the primary challenges is the scarcity of high-quality and standardized datasets, particularly for interfacial phenomena such as SEI formation, which are inherently complex and difficult to characterize. Moreover, many existing models fail to account for dynamic interactions between electrolyte components and electrode surfaces, leading to discrepancies between predicted and actual performance. This suggests that future research should focus on integrating ML with molecular simulations and experimental data to achieve more accurate and physically meaningful predictions.

4.3 Performance Prediction

Accurate prediction of battery performance is essential for evaluating the feasibility of new materials and guiding experimental efforts. Machine learning models have increasingly been used to predict key performance indicators, such as energy density, power density, rate capability, and cycle life, based on material properties and operational conditions. These predictive capabilities enable researchers to rapidly assess the potential of different material combinations without extensive experimental testing.

In sodium-ion battery research, performance prediction models often rely on a combination of material descriptors and cycling data to capture both intrinsic material properties and dynamic electrochemical behavior. Advanced models, including ensemble learning methods and deep neural networks, have demonstrated strong predictive performance in estimating cycle life and degradation trends. Time-series modeling approaches, such as recurrent neural networks, are particularly useful for analyzing cycling data and forecasting long-term battery behavior.

Despite these advancements, several limitations remain. Many performance prediction models are trained on datasets derived from lithium-ion batteries, which may not accurately reflect the unique characteristics of sodium-ion systems. This mismatch can lead to biased predictions and reduced model reliability. Additionally, the lack of standardized testing protocols for SIBs further complicates the development of generalizable models. Addressing these issues requires creating dedicated sodium-ion datasets and developing models that explicitly account for sodium-specific electrochemical phenomena.

4.4 Degradation and Lifetime Analysis

Understanding and mitigating degradation mechanisms is critical for improving the longevity and reliability of sodium-ion batteries. Machine learning has emerged as a valuable tool for analyzing complex degradation processes by leveraging large datasets obtained from cycling experiments, spectroscopy, and imaging techniques. These models can identify patterns and correlations that are difficult to detect with traditional analytical methods, enabling early prediction of battery failure and estimation of remaining useful life.

Recent research has focused on using ML models to detect early indicators of degradation, including changes in voltage profiles, capacity fading, and internal resistance growth. By analyzing these signals, ML can provide insights into underlying degradation mechanisms, including electrode structural collapse, electrolyte decomposition, and unstable SEI formation. Furthermore, data-driven approaches have been used to optimize operating conditions and charging protocols to minimize degradation and extend battery lifespan.

However, the application of ML in degradation analysis for SIBs is still in its infancy. One of the main challenges is the limited availability of long-term cycling data, which is essential for training reliable models. Additionally, degradation processes in sodium-ion batteries are often more complex and less understood than those in lithium-ion systems, making it difficult to develop accurate predictive models. Another important issue is the interpretability of ML results, as understanding the physical meaning of model predictions is crucial for translating insights into practical improvements. These challenges

underscore the need for more comprehensive datasets and for integrating ML with physics-based modeling approaches.

5. MATERIALS INFORMATICS FOR SIB

Materials informatics has emerged as a transformative paradigm in modern materials science by integrating data-driven approaches with physics-based modeling and experimental validation. In the context of sodium-ion batteries, materials informatics provides a systematic framework to accelerate the discovery and optimization of battery components by leveraging the synergy between density functional theory calculations, high-throughput simulations, experimental datasets, and machine learning algorithms. This integrated approach enables researchers to efficiently explore vast chemical and structural spaces that would otherwise be inaccessible to conventional experimental methods.

At the core of materials informatics for SIBs lies the concept of mapping complex structure–property relationships through data. In sodium-ion systems, these relationships are particularly intricate due to the strong coupling between ionic size effects, structural dynamics, and electrochemical behavior. By combining computational data with machine learning models, it becomes possible to identify non-obvious correlations between material descriptors and performance metrics, such as the influence of local atomic environments on sodium diffusion pathways or the role of crystal symmetry in stabilizing electrode structures during cycling. This capability significantly enhances the efficiency of materials screening and reduces the reliance on trial-and-error experimentation.

A typical materials informatics workflow for sodium-ion batteries begins with data acquisition, which may involve extracting information from experimental measurements, first-principles calculations, or publicly available materials databases. This is followed by data preprocessing and feature engineering, in which relevant descriptors are constructed to represent the compositional, structural, and electronic properties of materials. In SIB research, feature design often incorporates domain-specific knowledge, such as ionic radius mismatch, interlayer spacing, redox-active elements, and formation energy, which are critical for accurately capturing sodium storage behavior. The processed data are then used to train machine learning models to predict target properties or classify materials by their suitability for specific battery applications.

An important advancement in this field is the incorporation of high-throughput computational screening with active learning strategies. In such frameworks, machine learning models iteratively guide the selection of new candidate materials for simulation or experimental validation, thereby maximizing information gain while minimizing computational and experimental costs. This closed-loop approach is particularly beneficial for sodium-ion systems, where data scarcity remains a significant challenge. By selectively exploring the most informative regions of the materials space, active learning can accelerate the discovery of high-performance materials while continuously improving model accuracy.

Another critical aspect of materials informatics in SIB research is the integration of multi-scale data, ranging from atomic-level simulations to device-level performance metrics. This multi-scale perspective is essential for capturing the full complexity of battery systems, where phenomena occurring at different length and time scales are strongly interconnected. For example, atomic-scale diffusion barriers influence electrode kinetics, which in turn affect macroscopic performance indicators such as rate capability and cycle life. Machine learning models that span these scales offer significant potential to develop more accurate and holistic predictions of battery behavior.

Despite its promising potential, the implementation of materials informatics in sodium-ion battery research still faces several challenges. One of the primary limitations is the lack of standardized, high-quality datasets tailored to SIB systems. In addition, inconsistencies in data formats and measurement conditions can hinder the integration of datasets from different sources. Another challenge lies in the interpretability of machine learning models, particularly deep learning approaches, which often operate as black boxes. For materials informatics to be truly effective, it is essential to develop models that not only provide accurate predictions but also offer insights into the underlying physical mechanisms.

Looking forward, the future of materials informatics for sodium-ion batteries is expected to be shaped by the convergence of several emerging technologies. The integration of autonomous experimentation platforms, often referred to as self-driving laboratories, will enable rapid generation of high-quality datasets and real-time model validation. In parallel, advances in explainable artificial intelligence will enhance the transparency and reliability of machine learning models, facilitating their adoption in scientific research. Furthermore, the incorporation of quantum computing and quantum machine learning holds the potential to revolutionize materials discovery by enabling more accurate simulations of complex chemical systems. These developments collectively point toward a new era of data-driven battery research, where materials informatics serves as a central pillar in the design of next-generation sodium-ion energy storage technologies.

6. CHALLENGES AND LIMITATIONS

Despite the rapid advancement of machine learning applications in sodium-ion battery research, several fundamental challenges continue to hinder its full potential and broader adoption. One of the most critical limitations is the scarcity and inconsistency of high-quality datasets specifically tailored to sodium-ion systems. Unlike lithium-ion batteries, which benefit from extensive, relatively standardized datasets accumulated over decades of research, sodium-ion batteries still suffer from fragmented, heterogeneous data sources. Experimental results are often reported under varying conditions, with differences in material synthesis methods, testing protocols, and performance metrics, which makes it difficult to construct reliable and generalizable machine learning models. This lack of data standardization not only limits model accuracy but also complicates reproducibility and cross-study comparisons.

Another major challenge lies in the representation of materials and the selection of appropriate descriptors. In many cases, the performance of machine learning models is highly sensitive to feature engineering quality, particularly in systems as complex as sodium-ion batteries. While traditional descriptor-based approaches rely on handcrafted features derived from domain knowledge, they may fail to capture subtle structural and electronic interactions that significantly influence sodium storage behavior. On the other hand, advanced approaches such as graph neural networks aim to overcome this limitation by learning representations directly from atomic structures. However, these models often require large amounts of training data and substantial computational resources, which are not always available in SIB research. This creates a trade-off between model complexity and data availability that remains an open challenge.

Model interpretability is another critical issue that limits the practical utility of machine learning in sodium-ion battery development. While many models, particularly deep learning architectures, can achieve high predictive accuracy, they often operate as black boxes, offering little insight into the underlying physical mechanisms that govern battery performance. In scientific research, predictive capability alone is insufficient; understanding causality and deriving physically meaningful insights are equally important. The lack of interpretability can hinder the adoption of ML models by experimental researchers, who require clear guidance for material design and optimization. Although recent developments in explainable artificial intelligence have begun to address this issue, their application in sodium-ion battery research remains limited and warrants further exploration.

A further limitation is the issue of model transferability and generalization across different battery chemistries and material classes. Many machine learning models developed for battery research are trained on lithium-ion datasets due to their abundance, and are subsequently applied to sodium-ion systems. However, this approach often leads to suboptimal performance because of fundamental differences in ionic size, electrochemical potential, and structural dynamics between lithium and sodium systems. Even within sodium-ion batteries, models trained on specific classes of materials, such as layered oxides, may not generalize well to other systems, such as polyanionic compounds or Prussian blue analogs. This lack of transferability highlights the need for more diverse and representative datasets and for the development of domain-adaptive learning techniques.

In addition to data and modeling challenges, there is a significant gap in integrating machine learning with experimental workflows. While many studies demonstrate the predictive capabilities of ML models, relatively few successfully translate these predictions into experimentally validated materials or improved battery performance. This disconnect between computational predictions and practical implementation limits the real-world impact of machine learning in sodium-ion battery research. Bridging this gap requires closer collaboration between computational scientists and experimentalists, as well as the development of closed-loop systems that integrate prediction, synthesis, and characterization.

Finally, the dynamic and multi-scale nature of battery systems presents a fundamental challenge for machine learning approaches. Sodium-ion batteries involve complex interactions across multiple length and time scales, ranging from atomic-level diffusion processes to macroscopic performance behavior. Most existing ML models focus on a single scale, which limits their ability to capture the full complexity of battery operation. Developing multi-scale models that integrate information across different levels remains a significant research challenge, but also a key opportunity to advance the field.

These challenges collectively highlight that, while machine learning offers powerful tools for accelerating sodium-ion battery research, its current application is still in a relatively early stage. Addressing these limitations requires not only technical advancements in machine learning algorithms but also systematic efforts to improve data quality, standardization, and interdisciplinary collaboration. By overcoming these barriers, machine learning can play a more transformative role in the development of next-generation sodium-ion battery technologies.

7. FUTURE PERSPECTIVES

The future of machine learning in sodium-ion battery research is expected to evolve toward more integrated, autonomous, and physically informed frameworks that go beyond conventional data-driven prediction. One of the most promising directions is the development of closed-loop systems that combine machine learning models with high-throughput experimentation and automated synthesis platforms. In such self-driving laboratory environments, machine learning algorithms iteratively guide experimental design by selecting the most informative candidates, thereby accelerating the discovery of novel materials while minimizing experimental cost and time. For sodium-ion batteries, where data scarcity remains a critical bottleneck, this approach offers a practical pathway to continuously generate high-quality datasets and improve model reliability in real time.

Another key direction involves the advancement of physics-informed machine learning models that incorporate fundamental electrochemical principles into the learning process. Unlike purely data-driven approaches, these models integrate constraints derived from thermodynamics, kinetics, and solid-state physics, enhancing both predictive accuracy and interpretability. In sodium-ion systems, where complex interactions such as ion diffusion, phase transitions, and interfacial phenomena play a crucial role, embedding physical knowledge into ML models can significantly improve their ability to generalize across different material classes and operating conditions. This hybrid paradigm is particularly important for bridging the gap between computational predictions and experimental validation.

The integration of multi-scale modeling represents another important frontier in SIB research. Future machine learning frameworks are expected to simultaneously capture phenomena occurring at different length and time scales, from atomic-level ion transport to electrode-level structural evolution and full-cell performance. Such multi-scale approaches will enable a more comprehensive understanding of battery behavior and facilitate the design of materials and systems that are optimized across multiple performance metrics. Achieving this goal requires developing advanced data fusion techniques and hierarchical modeling strategies that can link diverse data sources.

In addition to methodological advancements, the expansion of open and standardized databases for sodium-ion batteries will play a crucial role in accelerating research progress. The establishment of community-driven data platforms with consistent formats, metadata standards, and benchmarking protocols will improve data accessibility, reproducibility, and model comparability. This is particularly important for addressing current limitations related to data fragmentation and inconsistency. Collaborative efforts among academia, industry, and research institutions will be essential to build such infrastructure and ensure its long-term sustainability.

Emerging paradigms such as transfer learning and active learning are also expected to significantly enhance the efficiency of machine learning in SIB research. Transfer learning enables the adaptation of knowledge from related domains, such as lithium-ion batteries, to sodium-ion batteries, thereby mitigating data scarcity. However, future research must focus on developing domain-aware transfer learning techniques that account for the fundamental differences between these chemistries. Meanwhile, active learning strategies can prioritize the acquisition of the most informative data points, further reducing the need for large datasets while improving model performance.

A particularly exciting frontier is the integration of quantum computing and quantum machine learning into sodium-ion battery research. Quantum computing has the potential to revolutionize materials science by enabling more accurate simulations of electronic structures and chemical interactions that are computationally prohibitive for classical methods. When combined with machine learning, quantum approaches can enhance the modeling of complex battery materials and provide deeper insights into sodium storage mechanisms at the quantum level. Although still in its early stages, this interdisciplinary field holds significant promise for overcoming current limitations in both computational accuracy and scalability.

Finally, the development of explainable artificial intelligence will be crucial for ensuring that machine learning models provide not only accurate predictions but also actionable scientific insights. Future research should focus on designing interpretable models that can reveal the underlying factors governing battery performance, such as identifying key structural features that influence ion diffusion or stability. This will facilitate more rational material design and strengthen the integration between computational and experimental research.

Overall, the future of machine learning in sodium-ion battery research lies in the convergence of data-driven methods, physical modeling, and experimental automation. By addressing current limitations and embracing emerging technologies, machine learning is poised to play a central role in the development of next-generation sodium-ion batteries that are both high-performing and economically sustainable.

8. CONCLUSION

This review has presented a comprehensive and critical overview of the role of machine learning in advancing sodium-ion battery research, highlighting its potential to transform traditional materials

discovery and battery optimization processes. By systematically examining the fundamental principles of sodium-ion batteries, various machine learning techniques, and their applications in electrode discovery, electrolyte design, performance prediction, and degradation analysis, this work underscores the growing importance of data-driven approaches in addressing the intrinsic limitations of sodium-based energy storage systems.

Despite significant progress, the application of machine learning in sodium-ion batteries remains relatively early compared to lithium-ion systems. Key challenges, including data scarcity, lack of standardization, limited model interpretability, and poor transferability, continue to hinder the development of robust and generalizable models. Furthermore, the gap between computational predictions and experimental validation underscores the need for more integrated, collaborative research frameworks.

This review contributes to the field by not only summarizing recent advancements but also critically identifying existing research gaps and outlining future directions. In particular, the integration of machine learning with materials informatics, high-throughput experimentation, and emerging paradigms such as quantum machine learning offers a promising pathway for accelerating innovation in sodium-ion battery technologies. The development of physics-informed and explainable models, along with the establishment of standardized and open-access datasets, will be essential for improving model reliability and fostering broader adoption within the scientific community.

In conclusion, machine learning holds significant promise as a key enabler for the next generation of sodium-ion batteries, providing the tools necessary to navigate complex material spaces and optimize performance with unprecedented efficiency. Continued advancements in this field, supported by interdisciplinary collaboration and technological innovation, are expected to play a pivotal role in realizing sustainable, cost-effective, and high-performance energy storage solutions.

REFERENCES

[1] Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. *Nature*, 559, 547–555.

[2] Ramprasad, R., Batra, R., Pilania, G., et al. (2017). Machine learning in materials informatics. *npj Computational Materials*, 3, 54.

[3] Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances of ML in materials science. *npj Computational Materials*, 5, 83.

[4] Jablonka, K. M., Ongari, D., Moosavi, S. M., & Smit, B. (2020). Big-data science in porous materials. *Nature Reviews Materials*, 5, 327–341.

[5] Unocic, R. R., et al. (2021). Data-driven design of battery materials. *Nature Reviews Materials*, 6, 723–743.

[6] Noé, F., Tkatchenko, A., Müller, K. R., & Clementi, C. (2020). Machine learning for molecular simulation. *Annual Review of Physical Chemistry*, 71, 361–390.

[7] Xu, Y., et al. (2020). Recent advances in sodium-ion batteries. *Advanced Energy Materials*, 10, 2001239.

[8] Wang, Y., et al. (2021). Sodium-ion batteries: materials and challenges. *Advanced Energy Materials*, 11, 2003155.

[9] Wu, X., et al. (2022). Challenges and perspectives of SIBs. *Nano Energy*, 95, 106944.

[10] Zhang, H., et al. (2023). Sodium-ion battery progress. *Energy Storage Materials*, 55, 123–145.

[11] Muhamad Akrom, Supriadi Rustad, Hermawan Kresno Dipojono, Ryo Maezono, Hideaki Kasai. Quantum machine learning for ABO3 perovskite structure prediction. Computational Materials Science, Volume 250, Pages 113694, 2025, https://doi.org/10.1016/j.commatsci.2025.113694.

[12] S Rustad, M Akrom, T Sutojo, HK Dipojono. A feature restoration for machine learning on anti-corrosion materials. Case Studies in Chemical and Environmental Engineering 10, 100902, 2024, https://doi.org/10.1016/j.cscee.2024.100902.

[13] Akrom, M., Rustad, S. & Dipojono, H.K. Investigation of Corrosion Inhibition Capability of Pyridazine Compounds via Ensemble Learning. J. of Materi Eng and Perform 34, 14948–14962 (2025). https://doi.org/10.1007/s11665-024-10129-x.

[14] Muhamad Akrom, Wise Herowati, De Rosal Ignatius Moses Setiadi. A quantum circuit learning-based investigation: A case study in iris benchmark dataset binary classification. Journal of Computing Theories and Applications. Volume 2, Issue 3, Pages 355-367, 2024, https://doi.org/10.62411/jcta.11779.

[15] Muhamad Akrom, Supriadi Rustad, Totok Sutojo, De Rosal Ignatius Moses Setiadi, Pulung Nurtantio Andono, Guruh Fajar Shidik, Hermawan Kresno Dipojono, Ryo Maezono. A novel quantum-enhanced model cascading approach based on support vector machine in blood-brain barrier permeability prediction. Materials Today Communications, Volume 45, Pages 112341, 2025, https://doi.org/10.1016/j.mtcomm.2025.112341.

[16] Muhamad Akrom, Usman Sudibyo, Achmad Wahid Kurniawan, Noor Ageng Setiyanto, Ayu Pertiwi, Aprilyani Nur Safitri, Novianto Hidayat, Harun Al Azies, Wise Herawati. Artificial Intelligence Berbasis QSPR Dalam Kajian Inhibitor Korosi. JoMMiT: Jurnal Multi Media dan IT, Volume 7, Issue 1, Pages 015-020, 2023, https://doi.org/10.46961/jommit.v7i1.721.

[17] M. Akrom, T. Sutojo, A. Pertiwi, S. Rustad, H.K. Dipojono, Investigation of Best QSPR-Based Machine Learning Model to Predict Corrosion Inhibition Performance of Pyridine-Quinoline Compounds, J Phys Conf Ser, 2673(1), 012014 (2023), https://doi.org/10.1088/1742-6596/2673/1/012014.

[18] M. Akrom, Green corrosion inhibitors for iron alloys: a comprehensive review of integrating data-driven forecasting, density functional theory simulations, and experimental investigation. J Mult Mater Inf, 1(1), 22–37 (2024), https://doi.org/10. 62411/jimat.v1i1.10495

[19] M. Akrom, S. Rustad, H.K. Dipojono, A machine learning approach to predict the efficiency of corrosion inhibition by natural product-based organic inhibitors, *Phys Scr*, 99(3), 036006 (2024), https://doi.org/10.1088/1402-4896/ad28a9.

[20] M. Akrom, S. Rustad, H.K. Dipojono, Machine learning investigation to predict corrosion inhibition capacity of new amino acid compounds as corrosion inhibitors, Results in Chemistry 6 (2023) 101126, https://doi.org/10.1016/j. rechem.2023.101126.

[21] M. Akrom, S. Rustad, A.G. Saputro, H.K. Dipojono, Data-driven investigation to model the corrosion inhibition efficiency of Pyrimidine-Pyrazole hybrid corrosion inhibitors, Comput. Theor. Chem. 1229 (2023) 114307, https://doi.org/10.1016/ J.COMPTC.2023.114307.

[22] M. Akrom, S. Rustad, H.K. Dipojono, Prediction of Anti-Corrosion performance of new triazole derivatives via Machine learning, Comput. Theor. Chem. 1236 (2024), https://doi.org/10.1016/j.comptc.2024.114599.

[23] M. Akrom, Investigation of natural extracts as green corrosion inhibitors in steel using density functional theory, *Jurnal Teori dan Aplikasi Fisika*, 10(1), 89-102 (2022), https://doi.org/10.23960%2Fjtaf.v10i1.2927.

[24] M. Akrom, S. Rustad, H.K. Dipojono. Development of quantum machine learning to evaluate the corrosion inhibition capability of pyrimidine compounds. Materials Today Communications, 39, 108758 (2024), https://doi.org/10.1016/j.mtcomm.2024.108758.

[25] M. Akrom, S. Rustad, H.K. Dipojono, SMILES-based machine learning enables the prediction of corrosion inhibition capacity, MRS Commun 14 (2024) 379–387, https://doi.org/10.1557/s43579-024-00551-6.

[26] M. Akrom, S. Rustad, A.G. Saputro, A. Ramelan, F. Fathurrahman, H.K. Dipojono, A combination of machine learning model and density functional theory method to predict corrosion inhibition performance of new diazine derivative compounds, Mater. Today Commun. 35 (2023) 106402, https://doi.org/10.1016/J. MTCOMM.2023.106402.

[27] M. Akrom, et al., DFT and microkinetic investigation of oxygen reduction reaction on corrosion inhibition mechanism of iron surface by Syzygium Aromaticum extract, Appl. Surf. Sci. 615 (2023), https://doi.org/10.1016/j. apsusc.2022.156319.

[28] M. Akrom, S. Rustad, H.K. Dipojono. Variational quantum circuit-based quantum machine learning approach for predicting corrosion inhibition efficiency of pyridine-quinoline compounds. Materials Today Quantum, 2, 100007 (2024), https://doi.org/10.1016/j.mtquan.2024.100007.

[29] Muhamad Akrom, Supriadi Rustad, Hermawan Kresno Dipojono, Ryo Maezono, Hideaki Kasai. Enhanced quantum support vector regression with quantum kernels and virtual sampling for ABX3 perovskite formation energy. Expert Systems with Applications, Pages 130817, 2025, https://doi.org/10.1016/j.eswa.2025.130817.

[30] Muhamad Akrom. Quantum Neural Network in Architectures, Learning Mechanisms, and Emerging Applications Across Domains: A Review. Journal of Multiscale Materials Informatics, Volume 2, Issue 2, Pages 30-39, 2025, 10.62411/jimat.v2i2.14929.