Journal of the Korean Electrochemical Society
Vol. 28, No. 4, 2025, 111-119
https://doi.org/10.5229/JKES.2025.28.4.111

Review Paper
pISSN 1229-1935
eISSN 2288-9000

머신러닝을 활용한 음이온교환막 기술: 데이터 표준, 표현 방식, 설명 가능성, 그리고 역방향 설계에 대한 리뷰

함지환[1] · 박진수[1,2]*

[1]상명대학교 공과대학 그린화학공학과
[2]상명대학교 미래환경×에너지연구소

(2025년 10월 27일 접수 : 2025년 11월 19일 수정 : 2025년 11월 19일 채택)

Machine Learning for Anion Exchange Membranes: A Review of Data Standards, Representations, Explainability, and Inverse Design

Ji-Hwan Ham[1] and Jin-Soo Park[1,2]*

[1]*Department of Green Chemical Engineering, College of Engineering, Sangmyung University, Cheonan 31066, Republic of Korea*
[2]*Future Environment and Energy Research Institute, Sangmyung University, Cheonan 31066, Republic of Korea*

(Received October 27, 2025 : Revised November 19, 2025 : Accepted November 19, 2025)

초 록

음이온교환막은 연료전지와 수전해 시스템의 핵심 구성 요소로, 이온전도도, 화학적 안정성, 치수 안정성, 그리고 기계적 강도 사이의 균형이 요구된다. 최근의 기계학습 연구들은 다양한 음이온교환막 화학 구조 전반에서 음이온전도도를 예측하고, 일부 경우에는 알칼리 안정성까지 예측할 수 있는 조건인지형 대리 모델을 제시하고 있다. 해석 가능한 분석 파이프라인은 고분자 구조와 시험 조건을 성능과 연결해주며, 비지도 매핑은 합성 이전 단계에서의 탐색을 지원한다. 또한 확장 가능한 역방향 설계 워크플로우는 기계 학습과 휴리스틱 탐색을 결합하여, 이온전도도–안정성 기준을 만족하는 거대한 가상 물질 라이브러리를 선별할 수 있게 한다. 장치 관점에서 보면, 데이터 기반 모델은 음이온교환막 수전해 성능을 지배하는 음이온교환막 특성과 운전 인자를 도출해낸다. 본 총설은 이러한 최근 발전을 정리하고, 조건 인지형 메타데이터, 다중 스케일 특성화, 해석 가능한 지도 학습, 비지도 매핑, 그리고 기계 학습 기반 스크리닝으로 이어지는 간결한 워크플로우를 제안함으로써 소재 수준에서 장치 수준으로 이어지는 재현 가능한 전환 경로를 제시한다.

Abstract : Anion exchange membranes (AEMs) underpin fuel cells and water electrolyzers, requiring a balance of ionic conductivity, chemical stability, dimensional control, and mechanical robustness. Recent machine learning (ML) studies introduce condition-aware surrogate models that predict anion conductivity and, in some cases, alkaline stability across diverse AEM chemistries. Interpretable pipelines relate structure and test conditions to performance, while unsupervised mapping supports pre-synthesis exploration. Scalable inverse-design workflows combine ML with heuristic search to screen large virtual libraries under conductivity-stability criteria. At the device level, data-driven models highlight membrane and operating parameters that govern anion exchange membrane water electrolysis (AEMWE) performance. This mini

*E-mail: energy@smu.ac.kr

review summarizes recent progress and outlines a concise workflow, i.e., condition-aware metadata, multi-scale featurization, interpretable supervision, unsupervised mapping, and ML-guided screening, for reproducible materials-to-device translation.

Keywords : Anion exchange membrane, Water electrolysis, Machine learning, Material-to-device translation

1. Introduction

In alkaline water electrolysis, AEM-based devices enable the use of non-noble catalysts while producing high purity hydrogen from renewable electricity. However, they still face ohmic losses and chemical instability of quaternary ammonium ionomers under strong alkali and elevated temperature.[1)]

Anion exchange membranes (AEMs) function as selective ionic conductors in electrochemical devices and thus must jointly deliver high ionic conductivity, chemical/alkaline durability, dimensional stability under hydration, and mechanical robustness during assembly and long-term operation.[2)] The long-recognized conductivity–dimensional stability tension especially constrains AEM design. In particular, increasing the ion exchange capacity (IEC) typically enhances ionic conductivity but also raises water uptake and swelling, which reduce dimensional and mechanical stability.[3)] Similarly, excessive hydration facilitates ion transport yet compromises structural robustness.[4)] Therefore, an optimal AEM should achieve a balanced optimization among all key properties including ionic conductivity, chemical durability, water uptake, and mechanical stability.[5)] However, in the early stage of AEM development, ionic conductivity has served as the most critical indicator for performance evaluation and still remains a major guiding metric in material design.[6)]

Traditional trial and error discovery struggles with the combinatorial space of backbones, cation distributions, side chain architectures, and processing histories. In addition, cross study reuse is hindered by inconsistent reporting of temperature, hydration mode (humidified vs. liquid fed), and alkaline aging conditions, which motivates condition-aware datasets and standardized analyses. Recent machine learning (ML) studies address these gaps by curating AEM datasets that explicitly encode test conditions and by training supervised models that predict hydroxide conductivity, including degraded states, with low error and human interpretable logic.[7)] Complementary deep learning protocols demonstrate accurate conductivity modeling for specific chemistries when representation and curation are handled carefully.[8)] Interpretable pipelines (e.g., shapley additive explanations, SHAP with descriptor models) and hybrid graph/descriptor frameworks connect structural and experimental variables to ionic conductivity, while unsupervised mapping reveals coverage and performance relevant clusters for pre-synthesis decisions across diverse chemistries.[9)] Expert augmented ML integrated with genetic algorithms has also operationalized multi-objective screening and down selection in large virtual copolymer libraries, supporting rapid iteration in AEM discovery.

In this mini-review, we synthesize these developments into a concise workflow, i.e., condition-aware metadata, multi-scale featurization, interpretable supervision with group-aware validation, unsupervised mapping, and ML-guided inverse design, and discuss their use from molecular design to device level optimization in membrane-based electrolyzers and fuel cells.

2. Standardized Data and Metadata for Machine Learning

Across AEM studies, measurement temperature, ion form (OH^- vs. Cl^- exchange), hydration mode (humidified vs liquid-fed), and alkaline aging protocols differ enough that pooling numbers without their conditions mixes fundamentally different response surfaces. As a result, models tend to learn condition-averaged artifacts instead of true structure–condition–property relationships.[7)] A practical solution is to record experimental conditions within each dataset and to separate fresh and degraded states so that reaction kinetics and degradation pathways are not confused. This practice has become standard

because it enables consistent modeling, prevents post-hoc guessing of test settings, and keeps extrapolation reliable when operating conditions change.[10)]

Data collection often begins with published tables and figures rather than raw repositories. Research teams usually digitize values, unify units and significant figures, and record uncertainty whenever possible. A minimal set of condition data typically includes three key factors time, temperature, and alkaline concentration along with measurement temperature, hydration mode, counter-ion, and sample history. Recording these details together with target variables allows the model to capture trends related to temperature or caustic strength instead of averaging them out.[8)]

Before modeling, corpus diagnostics and feature-ranking analyses identify which experimental and operational variables most affect model behavior. Fig. 1 shows the relative importance of major features in standardized AEM datasets, indicating that electrolyte concentration and temperature are the most dominant predictors of performance, followed by catalyst type and mechanical factors such as

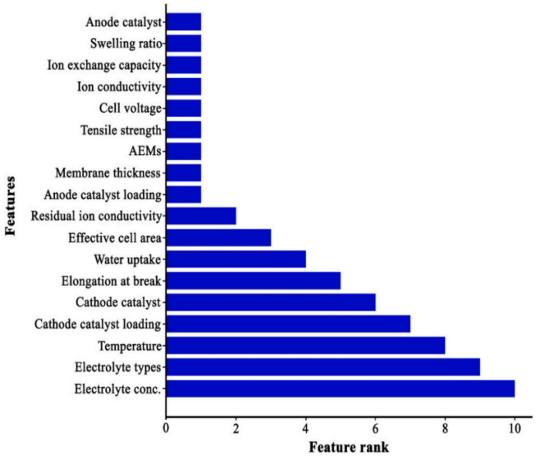

Fig. 1. Feature importance ranking for standardized AEM datasets obtained through recursive feature elimination (RFE) analysis. Key parameters such as electrolyte concentration, temperature, and catalyst loading appear as dominant predictors, emphasizing the importance of explicit metadata for condition-aware machine-learning models. Reproduced from Ref. 11, under the terms of the Creative Commons CC BY 4.0 license.

elongation at break and tensile strength.[11)] These rankings emphasize the importance of systematically recording and normalizing condition data to ensure reproducible machine learning studies.

At the device level, datasets from electrolyzers and fuel cells follow the same logic. Variables are grouped into three categories: general process parameters (current density, cell voltage, temperature, and feed strategy), membrane electrode assembly (MEA) parameters (electrode type, loading, and compression), and membrane properties (thickness, IEC, water uptake (WU), swelling ratio (SR), and mechanical strength). This organization allows researchers to compare data across laboratories using a single integrated table that combines both numerical and categorical data.[12)] For comparability, studies usually report the coefficient of determination (R^2) together with error metrics such as mean absolute error (MAE) and root mean square error (RMSE). Preprocessing steps including handling of missing values, outlier detection, and data scaling are also described. These standardized practices reduce bias, improve reproducibility, and make it possible to reanalyze data under different modeling approaches.[13)] These metadata and pre-processing conventions enable the subsequent learning steps, where multi-scale representations and supervised models translate raw measurements into structure-condition-property relationships.

3. Multi-Scale Representation and Explainable Supervised Models

Accuracy gains in AEM property prediction consistently trace back to condition-aware representation. Structural descriptors (backbone, substituents, composition, topology) combined with explicit test settings (temperature, hydration, ion form, aging flags) let models learn interactions rather than a single global trend.[7)] Hybrid pipelines that merge graph features (connectivity, ring systems, distances) with handcrafted descriptors tie structure and conditions to performance in ways that chemists can act on before synthesis. In practice, ranked drivers align well with human hypotheses and expose testable levers.[9)]

As feature spaces expand, we need parsimony

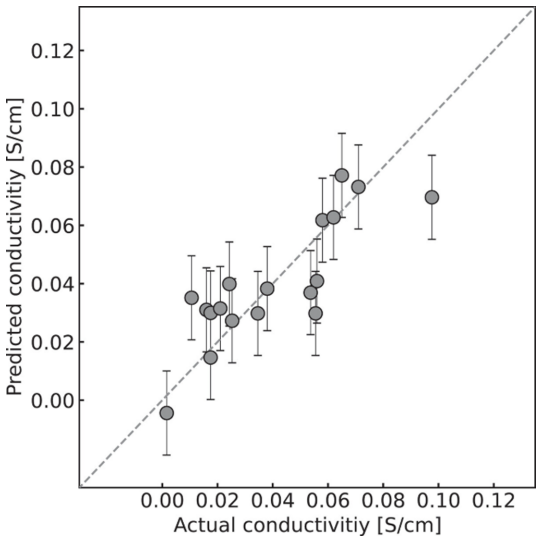

Fig. 2. Comparison between predicted and experimental hydroxide conductivity values from a Bayesian sparse-regression model demonstrating the model's predictive accuracy and consistency with experimental data. Reproduced from Ref. 14, under the terms of the Creative Commons CC BY 4.0 license.

without sacrificing fit. Bayesian sparse modeling retains performance while collapsing hundreds of candidates into a compact set that explains most variance. This yields readable 'design levers' with credible intervals, useful for down-selection and risk management.[14] As shown in Fig. 2, the Bayesian sparse regression model demonstrates a strong correlation between predicted and experimental hydroxide conductivity, highlighting its predictive robustness and interpretability. In particular, the model achieves an RMSE of approximately 0.014 $S \cdot cm^{-1}$, indicating high predictive reliability and supporting confidence in the trend visualized in Fig. 2.

SHAP is widely used for polymer ML models because it provides intuitive feature attributions that directly link structural variables to model decisions. To keep rationale attached to predictions, SHAP-based workflows quantify global and local attributions, identify interactions, and enable counterfactual analysis to determine which feature modifications would reverse or alter an outcome. This explainable framework can be broadly applied to other membrane properties, including ion conductivity and mechanical stability, and the resulting attributions can serve as priors for downstream searches.[15] CatBoost is particularly effective for polymer datasets because it can model nonlinear trends while naturally handling mixed numerical-categorical descriptors with minimal pre-processing. Operationally, teams stabilize comparisons with transparent baselines and reproducible training stacks. CatBoost often anchors tabular chemistry/processing tasks because it handles categorical variables, monotonic trends, and missingness gracefully, while PyTorch enables chemistry-aware networks and scalable training/inference when nonlinearity or sequence context adds signal.[16] Rather than ad-hoc tuning, hyperparameter optimization (HPO) such as Bayesian optimization (BO) and Hyperband standardizes model selection under limited data. BO efficiently searches hyperparameter spaces by prioritizing promising regions, whereas Hyperband accelerates tuning by adaptively allocating computational resources to the most promising configurations. Moreover, comparative evidence shows boosted trees can rival deep recurrent models on tabular targets, reinforcing a "baseline-first, escalate only as needed" ladder.[17]

These ingredients now appear in unified AEM playbooks that bundle representation choices, condition handling, and evaluation practices into a single recipe, and materials-informatics surveys provide the broader map so teams can adopt a shared language and avoid pipeline gaps.[1] That shared language sets the stage for unsupervised maps and validation protocols that stress-test what supervised models have actually learned.[18]

4. Data-Driven Insights: Unsupervised Mapping and Generalization Strategies

Unsupervised polymer maps built on curated structure plus both pristine-state temperature and the aging triplet visualize chemical coverage and performance-relevant clusters, flagging undersampled families, over-represented motifs, and high-performing pockets before a single new synthesis is commissioned.[10] Uniform Manifold Approximation and Projection (UMAP) is widely adopted because it preserves both local and global structure, enabling

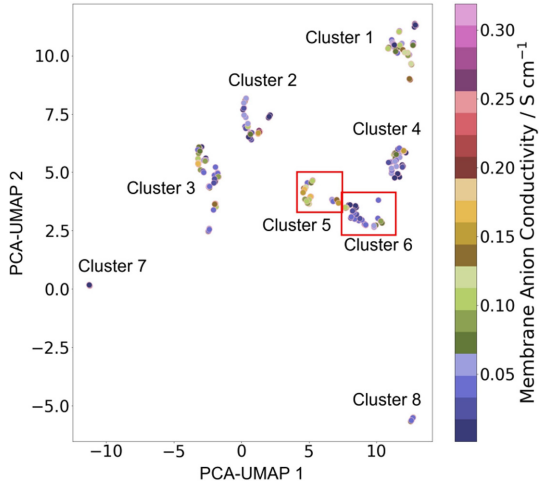

Fig. 3. Two-dimensional UMAP projection of anion-exchange-membrane (AEM) polymers based on molecular-descriptor embeddings. Each data point represents an individual polymer sample, colored by its experimentally measured hydroxide conductivity. Clusters (1–8) represent structurally or functionally similar subfamilies derived from k-means clustering. Low-density regions highlight under-explored chemical spaces that can guide future polymer design and screening efforts. Reproduced from Ref. 10, under the terms of the Creative Commons CC BY 4.0 license.

intuitive visualization of polymer families. As shown in Fig. 3, a two-dimensional UMAP embedding groups AEM polymers according to structural and conductivity similarity, forming eight clusters (Cluster 1–8) that represent distinct cationic backbone families and conductivity gradients. Low-density regions indicate unexplored chemical spaces that can guide targeted synthesis and model expansion, while spatial separation between clusters highlights how structural motifs and conductivity trends co-evolve. Such unsupervised embeddings also reveal areas where measurement conditions differ systematically, helping prevent the misinterpretation of temperature or hydration effects as purely structural causes during model interpretation.[9)]

Manifold projections such as Uniform Manifold Approximation and Projection (UMAP) and t-distributed Stochastic Neighbor Embedding (t-SNE) work together with clustering algorithms like Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Hierarchical DBSCAN (HDBSCAN). These methods complement supervised attributions by revealing structure–condition–performance patterns and identifying outliers that deserve targeted experiments. When feature importance values (for example, SHAP aggregates) are overlaid on the embeddings, they produce interpretable chemistry axes and condition sensitivities. These patterns can be used to define screening rules and set priorities for data augmentation.[19)] Model evaluation should also take the structure of the dataset into account. When polymer backbones or chemical families are repeated, random splits can make the model appear more accurate than it really is. Group-aware data splits and external validation are therefore needed to estimate generalization under realistic distribution shifts. Using consistent scenario-based protocols also enables reliable comparison across studies.[7)] This map-first, shift-aware approach connects data acquisition and modeling. It shows where to add data, which families to investigate, and how to divide datasets for fair validation. These insights directly support inverse-design workflows.[18)] Building on these map-driven insights, the next stage applies inverse-design strategies to explore virtual polymer spaces under competing conductivity-stability constraints.

5. Inverse Design and High-Throughput Screening

Insights from both supervised attributions and unsupervised structure maps naturally feed into inverse-design loops, enabling systematic exploration of vast copolymer design spaces. Genetic algorithms (GA) efficiently explore large copolymer libraries under multi-objective constraints. An expert-augmented ML + GA loop translates representation and attributions into action by exploring large virtual copolymer libraries under multi-objective criteria. It operates as a multi-objective loop combining ML and GA. It returns Pareto candidates that confront the trade-off among ionic conductivity, dimensional stability, and cation stability. It produces a short, defensible synthesis list with traceable levers for each entry.[20)] Because each candidate carries an explanation, lab teams can prioritize syntheses that probe model rationales rather than only predicted

scores.[15]

Scaling this loop demands explainability that survives throughput and infrastructure that keeps iteration times down. SHAP-assisted design reuses attributions as priors for GA/BO to shrink search while keeping the explicitly attached rationale on why each candidate is selected. On the engineering side, tabular baselines remain the audit-ready backbone, and custom neural models are slotted in only where nonlinearity buys measurable lift.[21] To ensure fairness across libraries and feature sets, model choice and tuning follow reproducible HPO procedures rather than manual tweaking, which also clarifies when variance comes from data versus search.[22]

Morphology–transport evidence sharpens feasibility filters. Experimentally demonstrated continuous ion-transport pathways and well-connected conducting channels justify cutoffs that align virtual down selection with what can actually be built and tested.[23] On the stability axis, polyaromatic ionomer case studies supply alkaline-durability bounds and failure modes that act as guardrails during down-selection and MEA integration, accelerating the candidate to prototype handoff.[24] These candidate-level insights ultimately require validation and optimization at the device level, where membrane properties interact with operating conditions to shape AEMWE performance.

Recent AEM informatics studies demonstrate that ML-guided inverse design is rapidly becoming a central methodology rather than an optional extension. Large-scale workflows that couple supervised surrogates with genetic algorithms have successfully screened millions of hypothetical copolymers and returned experimentally validated, Pareto-optimal candidates that balance hydroxide conductivity, swelling-driven dimensional stability, and alkaline durability.[20] In parallel, SHAP and Bayesian derived structural levers provide chemically interpretable priors such as the influence of backbone rigidity, cationic substitution patterns, and hydrophobic-hydrophilic domain tuning that guide rational modification before synthesis.[15] These advances collectively show that inverse design now offers a practical and scalable route for navigating vast polymer spaces, accelerating the transition from empirical discovery toward data-driven molecular engineering.

Finally, broader design reviews codify the representation-surrogate-search grammar and elevate uncertainty and reproducibility to first-class concerns so screening results remain comparable and cumulative across campaigns and labs, especially once device-level operating windows are folded into the objective.[12]

6. From Molecular Design to Device-Level Optimization

At the device level, data-driven AEMWE analyses decompose the relative influence of membrane descriptors and operating variables, yielding importance rankings and partial-dependence trends based on explicit temperature, feed, and current-density settings.[25] Adding knowledge-integrated frameworks embeds physics priors and domain rules in the learning loop, improving extrapolation when design spaces move beyond the training corpus and stabilizing recommendations across operating scenarios.[26]

Prediction-guided optimization then visualizes trade-offs among current density, temperature, and feed strategy, proposing operating windows that can be validated in staged tests. As shown in Fig. 4, CatBoost-based partial-dependence analyses reveal how anode catalyst loading, AEM type, ionic conductivity, and cell voltage each shape the current-density response, outlining monotonic and saturation behaviors that define efficient operating regimes for AEMWE optimization. Importantly, the CatBoost model supporting these trends exhibits strong predictive accuracy, achieving $R^2 \approx 0.90$ with MAE around 0.02, ensuring that the parameter–performance relationships reflect robust model-learned behavior. Because membrane-level descriptors, i.e., IEC, WU, SR, thickness, and pristine/aged conductivity are passed as first-class inputs, stack-level losses, efficiency, and durability can be traced back to specific levers rather than treated as black-box phenomena.[11] In parallel, polyaromatic-ionomer case studies delineate durability thresholds under strong alkali, offering directly actionable baselines for down-selection and MEA integration that keep device models grounded in materials reality.[24]

For deployment and monitoring, evaluation pairs

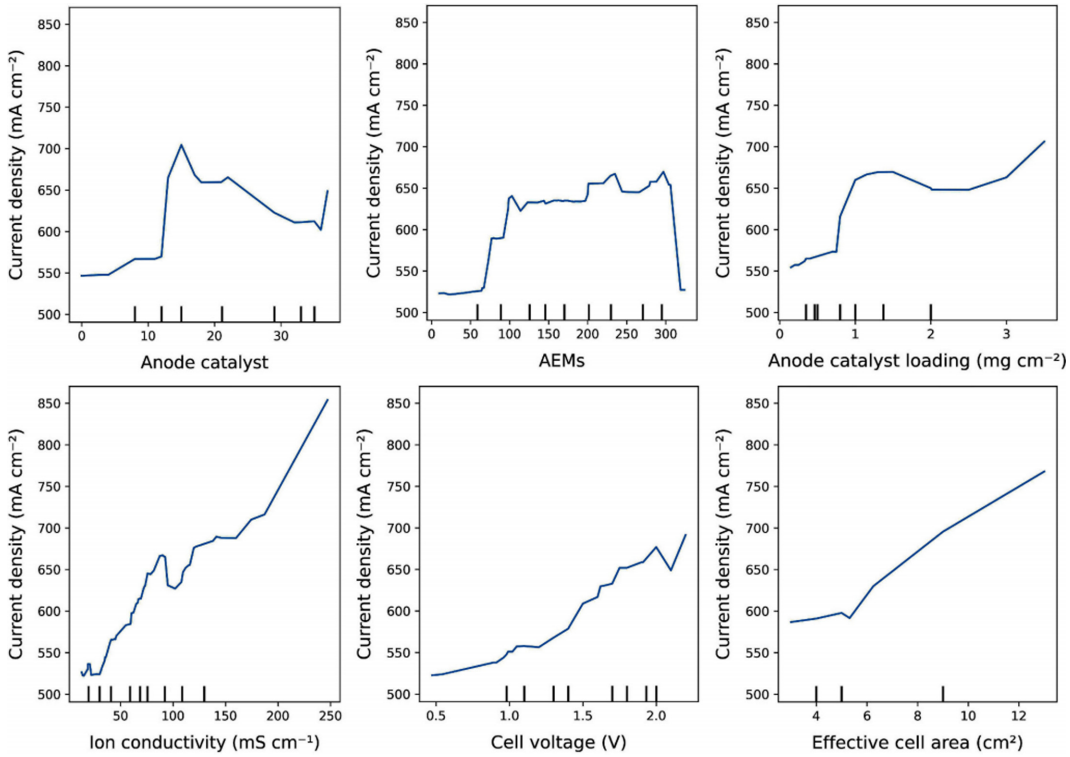

Fig. 4. Partial-dependence plots (PDPs) showing the marginal effects of key operating variables, i.e., anode catalyst, anion-exchange-membrane (AEM) type, catalyst loading, ionic conductivity, cell voltage, and effective cell area, on the predicted current density of AEM water-electrolysis systems. The trends delineate efficient operating windows that balance catalyst utilization and ion-transport enhancement. Reproduced from Ref. 11, under the terms of the Creative Commons CC BY 4.0 license.

R^2 with error metrics (MAE/RMSE) and fixes reproducible HPO procedures so tuning variance is controlled across regimes and the handoff from materials-level surrogates to stack-level targets is smooth.[13] Materials-informatics surveys systematize the bridge from representation to deployment, while anion exchange membrane fuel cell (AEMFC) practice summarizes membrane chemistries, fuels/electrolytes, and durability constraints that anchor electrolyzer selection in realistic operating envelopes. Together, they close the materials device loop without sacrificing interpretability.[22]

Practical guidance from monitoring/estimation benchmarks such as comparative state-of-charge (SoC) estimation studies on tabular targets reinforces transparent model/metric choices and scenario-aware validation, which are the same habits that keep electrolyzer optimization robust as conditions shift.[17]

7. Limitations and Threats to Validity

Despite rapid progress in machine learning-assisted design of AEMs, several factors continue to constrain the reliability and generalization of existing approaches. Most available datasets are biased toward quaternary ammonium backbones and hydroxide ion forms, limiting extrapolation to alternative cationic chemistries or mixed ion systems and narrowing the diversity of accessible design space.[8,11] Furthermore, publication bias remains a persistent issue, as the literature disproportionately reports high performing or stable membranes while neglecting failed syntheses and degraded samples, resulting in overly optimistic baselines and incomplete uncertainty characterization.[9] Models trained within narrow windows of temperature, humidity, or alkaline concentration also tend to exhibit instability

when applied under untested conditions. Although multi-condition frameworks that embed temperature and caustic strength as metadata partially mitigate this issue, their extrapolation capacity remains bound to the underlying data distribution.[14] Another recurring limitation involves data leakage: when polymer variants with similar backbones appear across training and validation folds, apparent model accuracy can be inflated. Group-aware cross-validation and backbone-level splitting have therefore been proposed to ensure statistical independence and preserve true generalization.[1]

Finally, model complexity often introduces a trade-off between predictive power and interpretability. Deep learning or ensemble-based architectures can achieve higher R^2 and lower MAE but obscure mechanistic relationships, making it difficult to derive chemical insight. Bayesian-sparse or SHAP-assisted frameworks have been suggested as intermediate solutions, balancing accuracy with physical transparency by assigning credible intervals and feature attributions to the learned relationships.[15,22]

Moving forward, reproducible AEM informatics will require community-level benchmark datasets with explicit uncertainty labels, transparent inclusion of degradation and failure data, and standardized validation protocols that align statistical rigor with physicochemical interpretability. Such collective practices will strengthen generalization, reduce optimism bias, and support more reliable translation of machine learning insights into actionable AEM design.

8. Concluding Remarks

ML now provides a practical advantage for AEM-based electrolyzers by turning scattered measurements into condition-aware, testable guidance. With standardized metadata, multi-scale representations, and explainable models, teams can down-select polymers faster, map operating windows, and link membrane levers to stack performance without guesswork. Unsupervised maps expose gaps before synthesis, and inverse design loops convert attributions into defensible candidate lists that respect conductivity-stability trade-offs. What remains is execution at scale by expanding condition-aware corpora, fixing shared benchmarks, and coupling fresh and aged states in one workflow. Maintaining a lab-in-the-loop workflow ensures that model recommendations are rapidly validated. AEMs must achieve both high performance and durability under strong alkali, necessitating the development of durable ionomers. In addition, robust device tuning should advance in parallel. This developed playbook shortens the path from data to deployment and moves AEMWE closer to commercially credible operation.

Acknowledgments

This research was supported by Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ016253), Rural Development Administration, Republic of Korea.

References

1. Y. Yuan, P. Fang, H. Yuan, X. Zou, Z. Sun, and F. Yan, Machine learning for prediction and synthesis of anion exchange membranes, *Acc. Mater. Res*., **6(3)**, 352–365 (2025).
2. J.-H. Park, Y. Park, T.-S. Jeon, Y. Seo, and J.-S. Park, Impact of cross-linking-monomer characteristics on pore-filling-membrane performance and durability in anion-exchange water electrolysis, *Appl. Sci.*, **15(3)**, 1495 (2025).
3. J.-S. Park, Toward scalable and robust diffusion-osmotic membranes for sustainable salinity gradient energy harvesting, *Appl. Chem. Eng.*, **36(4)**, 381–390 (2025).
4. D. Jeong and J.-S. Park, Effect of anion-conducting electrolytes in pore-filling membranes on performance and durability in water electrolysis, *Membranes*, **14(12)**, 265 (2024).
5. M. Lee and J.-S. Park, Enhanced performance and durability of pore-filling membranes for anion-exchange membrane water electrolysis, *Membranes*, **14(12)**, 269 (2024).
6. S. Bae, J. Park, S. B. Bong, J.-S. Park, B. Jeong, and J. Lee, Pore surface engineering of FeNC for outstanding power density of alkaline hydrazine fuel cells, *Chem. Eng. J.*, **479**, 147522 (2024).
7. Y. K. Phua, T. Fujigaya, and K. Kato, Predicting the anion conductivities and alkaline stabilities of anion-conducting polymers, *Sci. Technol. Adv. Mater.*, **24(1)**, 2261833 (2023).
8. F.-H. Zhai, Q.-Q. Zhan, Y.-F. Yang, N.-Y. Ye, R.-Y. Wan, J. Wang, S. Chen, R.-H. He, A deep learning protocol for

analyzing and predicting ionic conductivity of anion exchange membranes, *J. Membr. Sci.*, **642**, 119983 (2022).
9. P. Naghshnejad, J. A. Romagnoli, R. Kumar, and J. J. Chen, Uncovering structure–conductivity relationships in anion-exchange membranes using interpretable machine learning, *ChemRxiv* (2025).
10. Y. K. Phua, N. Terasoba, M. Tanaka, T. Fujigaya, and K. Kato, Unsupervised machine learning-derived anion-exchange membrane polymers map: a guideline for polymers exploration and design, *ChemElectroChem.*, **11**, e202400252 (2024).
11. M. M. Kabir, Y. Choden, S. Phuntsho, L. Tijing, H. K. Shon, Predictive machine learning optimization of anion exchange membrane water electrolysis systems, *Desalination*, **614**, 119198 (2025).
12. Z. Cao, O. B. Farimani, J. Ock, and A. B. Farimani, Machine learning in membrane design: from property prediction to AI-guided optimization, *Nano Lett.*, **24(10)**, 2953–2960 (2024).
13. D. Chicco and G. Jurman, The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression evaluation, *PeerJ Comput. Sci.*, **7**, e623 (2021).
14. R. Murakami, K. Miyatake, A. M. A. Mahmoud, H. Yoshikawa, and K. Nagata, Bayesian sparse modeling for interpretable prediction of hydroxide ion conductivity in anion-conductive polymer membranes, *Sci. Technol. Adv. Material, Meth.*, **5(1)**, 2534320 (2025).
15. R. J. Dalal, F. Oviedo, M. C. Leyden, and T. M. Reineke, Polymer design via SHAP and Bayesian machine learning optimizes pDNA and CRISPR ribonucleoprotein delivery, *Chem. Sci.*, **15**, 7219–7228 (2024).
16. N. Safaei, B. Safaei, S. Seyedekrami, M. Talafidaryani, A. Masoud, S. Wang, Q. Li, and M. Moqri, E-CatBoost: An efficient machine learning framework for predicting ICU mortality using the eICU collaborative research database, *PLoS One*, **17(5)**, e0262895 (2022).
17. A. Zilali, M. Adda, K. Ziane, and M. Berger, Machine learning-based state-of-charge estimation: a comparison between CatBoost and C-BLSTM-AE models, *Mach. Learn. Appl.*, **20**, 100629 (2025).
18. R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, Machine learning in materials informatics: recent applications and prospects, *npj Comput. Mater.*, **3**, 54 (2017).
19. Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature*, **521**, 436–444 (2015).
20. L. Liu, Y. Li, J. Zheng, and H. Li, Expert-augmented machine learning to accelerate the discovery of copolymers for anion exchange membranes, *J. Membr. Sci.*, **693**, 122327 (2024).
21. A. Paszke, S. Gross, F. Massa, *et al.*, PyTorch: an imperative style, high-performance deep learning library, *Adv. Neural Inf. Process. Syst.*, **32**, 8026–8037 (2019).
22. L. Yang and A. Shami, On hyperparameter optimization of machine learning algorithms: theory and practice, *Neurocomputing*, **415**, 295–316 (2020).
23. Y. Li, J. Zhang, H. Yang, S. Yang, S. Lu, H. Wei, and Y. Ding, Boosting the performance of an anion exchange membrane by the formation of well-connected ion conducting channels, *Polymer Chem.*, **10**, 2822–2831 (2019).
24. S. Maurya, S. Noh, I. Matanovic, E. J. Park, C. N. Villarrubia, U. Martinez, J. Han, C. Bae, and Y. S. Kim, Rational design of polyaromatic ionomers for alkaline membrane fuel cells with >1 W cm^{-2} power density, *Energy Environ. Sci.*, **11**, 3283–3291 (2018).
25. R. Ding, Y. Chen, Z. Rui, K. Hua, Y. Wu, X. Li, X. Duan, X. Wang, J. Li, and J. Liu, Guiding the optimization of membrane electrode assembly in a proton exchange membrane water electrolyzer by machine learning modeling and black-box interpretation, *ACS Sustain. Chem. Eng.*, **10**, 4561–4578 (2022).
26. X. Chen, A. Rex, J. Woelke, C. Eckert, B. Bensmann, R. Hanke-Rauschenbach, and P. Geyer, Machine learning in proton exchange membrane water electrolysis — A knowledge-integrated framework, *Appl. Energy*, **371**, 123550 (2024).