

Machine Learning Optimization of Energy Materials Design and Development

K. Daniel^{1,*}, Shaik Azad Basha^{2,3}, K. Srinivasarao⁴, D. Bhadra Rao⁵, CH. Naga Mani⁶, Syed Khasim^{7,*}, G. Megala⁸

¹Department of Physics, Dr. Samuel George Institute of Engineering and Technology, Markapur-523316, India.

²Department of Physics, Acharya Nagarjuna University, Guntur-522510, Andhra Pradesh, India.

³Department of Biotechnology, Alliance School of Applied Engineering, Alliance University, Bengaluru - 562106, Karnataka, India.

⁴Department of Physics, Government Degree College for Men, Srikakulam-532001, A.P, India

⁵Department of Physics, Adithya University, Surampalem, Kakinada-533437, A.P, India

⁶Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, A. P, India

⁷Department of Computer Science & Engineering, Dr. Samuel George Institute of Engineering and Technology, Markapur-523316, India.

⁸School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore - 632014, Tamil Nadu, India.

*Corresponding Author: Kodamala Daniel: danikodamala@gmail.com, Syed Khasim: khasim716@gmail.com
Orcid ID: 0009-0003-8525-122X, 0000-0003-1967-5449,

Abstract

Machine learning (ML) is quickly becoming a core tool for speeding up the discovery as well as the design and development of various materials used in energy conversion and storage. The joint use of data-driven models, density functional theory (DFT), high-throughput experimentation and material design approaches have enabled scientists to explore and screen the chemical space that were computationally or experimentally inaccessible. This article describes most commonly used ML techniques for the study of energy materials, presents major applications in areas of electrolytes and electrodes batteries, photoelectric materials, electrocatalysts and thermoelectrics and discusses the main issues such as small quantity of data at disposal, lack of understanding and low performance prediction that have limited the implementation of these materials in real life.

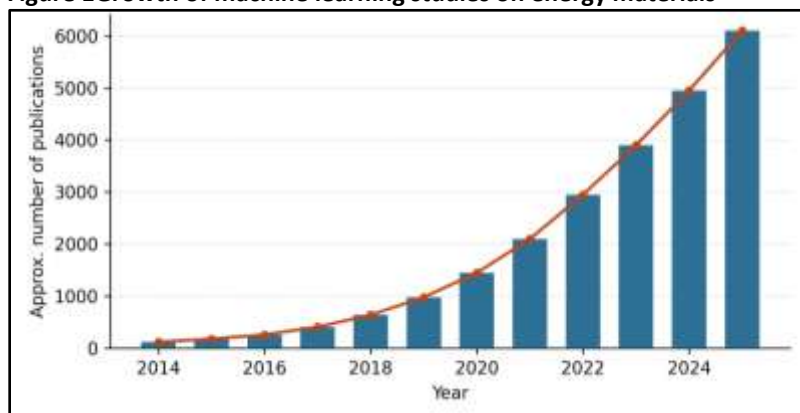
Keywords: Machine learning, Energy materials, Battery materials, photovoltaics, Graph neural networks.

1. Introduction

Worldwide, the only way for moving the world away from fossil fuels toward clean and low-carbon energy sources is through new materials that can store, convert or deliver energy more efficiently than our existing ones[1]. But the search for such materials via a purely experimental approach or even the first-principles simulation alone is very time-consuming, there is a chemical and compositional space of candidate materials that is so huge that no person can imagine it, while the DFT calculations and synthesis campaigns are still computationally and experimentally very expensive[2,3]. Machine learning offers an alternative way of dealing with such a dilemma, the system learns the pattern from the material descriptors and target properties of the existing materials and then it uses the obtained pattern to screen, rank or even create new materials[4].

The interdisciplinary area of machine learning and energy-materials research has garnered significant attention in the last ten years, as shown graphically in **Figure 1**. This graph captures the general increase of publications which combine both domains in the fields of battery, photovoltaic and catalysis. The present review paper summarizes the topic by presenting a compact and applied outline of ML techniques and how they work together in different energy material systems[5,6][7].

Figure 1 Growth of machine learning studies on energy materials



2. Machine-Learning Methods for Materials Discovery

2.1 Data and Descriptors

Machine learning-based materials depend on features that translate composition, crystal structure, or processing conditions into numbers. Feature categories are typically compositional/stability descriptors, crystal graph-based structural fingerprints and Spectro electrochemical measurements. The training sets are from computational databases, carefully compiled experimental sets of data and nowadays also from text-mined scientific documents[8,9].

2.2 Model Classes

Different algorithm families have been explored in energy materials, ranging from simple linear models to graph neural networks that work directly with atomic structures. **Table 1** contains a summary of the main families, their usual applications and an overview of the strengths and limitations[10–12].

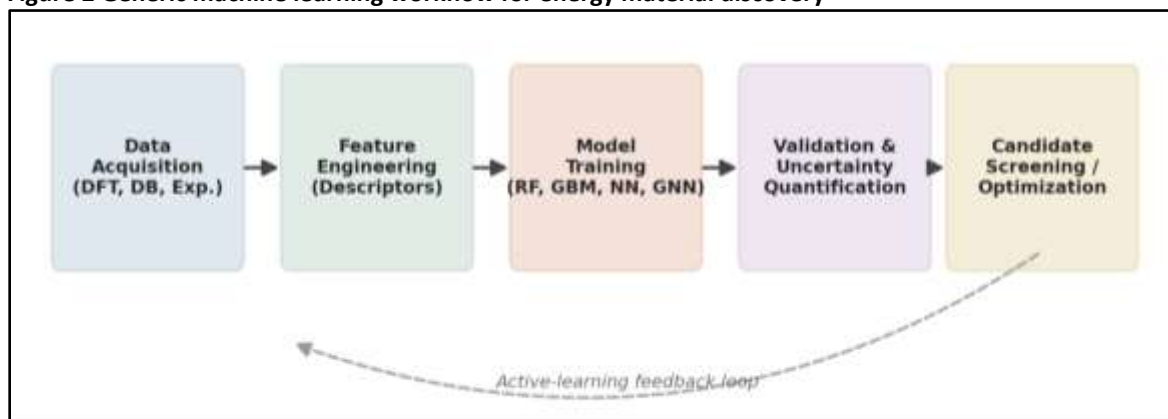
Table 1 Summary of machine-learning algorithm families commonly applied to energy-materials research

Algorithm family	Typical use case	Strengths	Limitations
Linear / LASSO / Ridge	Composition-property screening, feature selection	Interpretable, fast, low data requirement	Cannot capture strong nonlinearity
Random Forest / Gradient Boosting	Property prediction from tabular descriptors (e.g., XRD, XPS)	Robust, handles mixed features, feature importance	Limited extrapolation beyond training domain
Gaussian Process / Bayesian Optimization	Small-data optimization (bandgap tuning, synthesis conditions)	Built-in uncertainty, sample-efficient	Scales poorly with dataset size / dimensionality
Deep Neural Networks	Large-scale property/spectra regression	High capacity, flexible architectures	Data-hungry, limited interpretability
Graph Neural Networks	Crystal/molecular structure-property learning	Learn directly from atomic structure; strong accuracy	Computationally heavier; needs structural data

2.3 General ML Workflow

Regardless of the differing approaches, the bulk of machine learning research in this area typically adheres to a similar workflow obtaining data, generating descriptors/features, developing the model, evaluating its performance while quantifying uncertainties and finally performing candidate screening or optimization[13]. Such a method may be enhanced with an active-learning feedback loop that identifies the most informative upcoming experiment or calculation to be carried out next (**Figure 2**). Mainly in the case that experiments are costly and/or simulations require high-fidelity or only very limited data are available, Bayesian optimization and techniques based on Gaussian processes are the choices[14–16].

Figure 2 Generic machine learning workflow for energy material discovery



3. Energy Materials day-to-day applications

3.1 Battery Materials

Several examples from the field of battery research show different areas wherein ML has found its applications like screening cathode/anode compositions, designing electrolyte and predicting degradation and lifetime. The authors combined different methods like Gaussian process and DFT to search for solid-state electrolytes that are both mechanically strong and conduct ions highly simultaneously[17]. Meanwhile, ML with good explanatory powers together with a ML strategy that adjusts itself to situations have given an ability to predict cycle life under various working conditions and have provided means of battery health management and second-life sorting for purposes of recycling[18–20].

3.2 Photovoltaic Materials

One successful example is identifying potential materials for solar-cell absorbers. A team combined high-throughput DFT calculations with gradient-boosting and classification modeling to discover and select the best halide/perovskite mixed-composition compounds for a nearly ideal single-junction bandgap[21]. They achieved around a 89 % classification accuracy (ROC-AUC) and a 92 % regression accuracy (R), as representative studies. Such automated pipelines make it possible to computationally screen a few thousand different materials and thereby rank them [22–24].

3.3 Electrocatalysts

The design of an electrocatalyst mainly relies on atomic-level changes through machine learning. The active areas that can't be viewed easily makes this method all the more appealing as a means of understanding how atomic arrangement and electron configuration of the catalyst influence its activity[25]. The limited set of training data have not prevented multi-view ML setups from giving a reasonably accurate account of electrocatalytic sites within lithium sulphur battery materials, and the larger picture of ML in electrocatalysis, for example, shows the application of random forests and neural networks in interpreting XPS/XRD characterization data and finding catalysts that break up the so-called scaling laws at the same time[26–28].

3.4 Thermoelectric and Other Functional Materials

Other than batteries, photovoltaics and catalysts, ML can also be a great support in finding materials of other types like thermoelectric, piezoelectric and superconducting through structure-property relationships that are learned from composition and structural descriptors. The graph neural network models like CGCNN and SchNet were initially conceived for a general materials property prediction task but their application [29]. **Figure 3** shows the typical performance levels that have been reported across these four disciplines. **Table 2** gives examples of studies, methods and outcomes that are mainly useful[30].

Figure 3 Representative ML performance across energy material domains

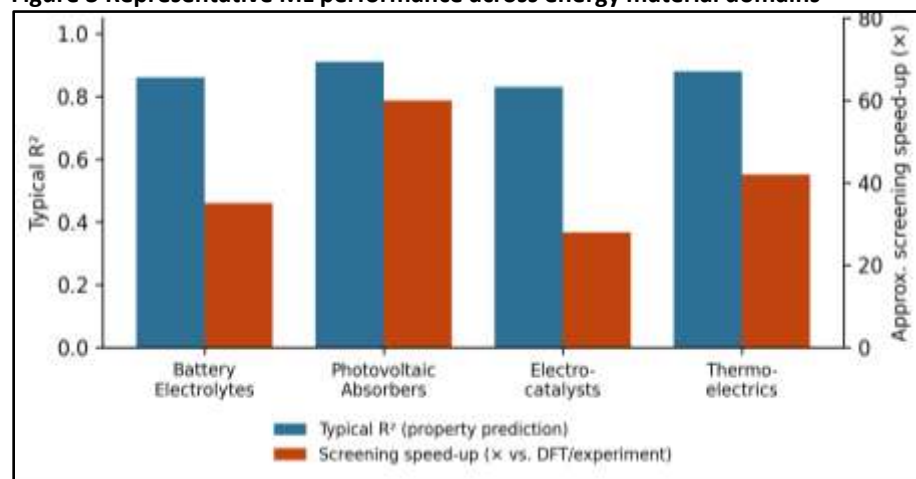


Table 2 Representative machine-learning studies across major energy-material application domains

Domain	ML approach	Task	Key reported outcome
Battery electrolytes	Gaussian Process + DFT	Solid-state electrolyte ionic-conductivity screening	Accelerated identification of mechanically stable, high-conductivity candidates
Battery lifetime	Interpretable feature-engineered ML	Cycle-life / degradation prediction across operating conditions	Cross-condition lifetime prediction with physically interpretable features
Photovoltaic absorbers	Gradient boosting + high-throughput DFT	Screening of halide-perovskite and mixed-halide absorbers	ROC-AUC ~0.89 classification; R^2 ~0.92 regression for optimal bandgap candidates
Electrocatalysts	Multi-view ML on limited datasets	Active-site evaluation for Li-S electrocatalysts	Improved mechanistic understanding of catalytic sites from sparse data
General materials informatics	Graph neural networks (CGCNN)	Crystal structure to property mapping	Accurate, interpretable prediction directly from crystal graphs

4. Conclusion

Machine learning has now become an essential part of the energy-materials discovery pipeline and it helps to screen faster various battery photovoltaic catalytic and thermoelectric materials with greater efficiency than conventional computational or experimental techniques would achieve alone. And, progress will hinge on tackling long-standing challenges over data availability, interpretability and experiment validation, while then again, the novel techniques which include foundation models, physics-informed architectures and autonomous experimentation will likely play roles of great impact in accelerating energy materials discovery at large.

References

- [1] ur Rehman A, Sanjari MJ, Elavarasan RM, Jamal T. Sustainability-aligned pathways for energy transition: A review of low-carbon energy network solutions. *Renewable and Sustainable Energy Reviews* 2026;226:116428. <https://doi.org/10.1016/J.RSER.2025.116428>.
- [2] Kaur M, Randhawa P, Jaiswal J, Dubal D, Bulakhe RN, Balakrishnan D, et al. Data-Driven Materials Science Using Machine Learning and Computational Modeling. *Computers, Materials and Continua* 2026;88. <https://doi.org/10.32604/CMC.2026.079503>.
- [3] Basha SA, Cole S, Divakaran D, Suyambulingam I. Solid State Reaction Method for Nanomaterials Synthesis: A Comprehensive Review on Characterization, Properties, and Applications. *Springer Proceedings in Materials*, vol. 60, Springer; 2024, p. 667–89. https://doi.org/10.1007/978-981-97-7071-7_47.
- [4] Guo H, Li S, Li S. Applications of Machine Learning in Polymer Materials: Property Prediction, Material Design, and Systematic Processes. *Computers, Materials and Continua* 2026;87. <https://doi.org/10.32604/CMC.2026.076492>.
- [5] Mobarak MH, Mimona MA, Islam MA, Hossain N, Zohura FT, Imtiaz I, et al. Scope of machine learning in materials research —A review. *Applied Surface Science Advances* 2023;18:100523. <https://doi.org/10.1016/J.APSADV.2023.100523>.
- [6] Divakaran D, Sriariyanun M, Basha SA, Suyambulingam I, Sanjay MR, Siengchin S. Physico-chemical, thermal, and morphological characterization of biomass-based novel microcrystalline cellulose from *Nelumbo nucifera* leaf: Biomass to biomaterial approach. *Biomass Convers Biorefin* 2023. <https://doi.org/10.1007/s13399-023-04349-9>.
- [7] França RP, Borges Monteiro AC, Arthur R, Iano Y. An overview of deep learning in big data, image, and signal processing in the modern digital age. *Trends in Deep Learning Methodologies: Algorithms, Applications, and Systems* 2020:63–87. <https://doi.org/10.1016/B978-0-12-822226-3.00003-9>.

- [8] Datta S, Chakraborty R, Bhattacharjee R, Kar AK. A review of artificial intelligence in materials screening: Data, models, and applications in forensic science. *Next Materials* 2026;13:102756. <https://doi.org/10.1016/J.NXMATE.2026.102756>.
- [9] K.Daniel, K.Venkatarao, K.Srinivasarao, Shaik Azad Basha, Sandhya Cole. Structural and Optical bandgap studies of undoped and Sm³⁺ doped MgZn₂(PO₄)₂ nanopowders. *ANU Journal of Physical Sciences* 2021;13:53–60.
- [10] Ekengwu IE, Ekengwu BO. Machine learning-accelerated inverse design of energy materials: A critical review of graph neural networks, physics-informed models, and generative AI for batteries, perovskite solar cells, and electrocatalysts. *Next Materials* 2026;13:102969. <https://doi.org/10.1016/J.NXMATE.2026.102969>.
- [11] Balasubramanyam SV, Rao KS, Daniel K, Basha SA, Cole S. Structural and Optical Insights of Sm³⁺ Doped MgCdP₂O₇ Nanopowders for Solid-State Lighting Applications. *Physical Chemistry Research* 2025;13:463–78. <https://doi.org/10.22036/PCR.2025.507953.2645>.
- [12] Rani GR, Kuna S, Kodamala D, Basha SA, Cole S. Structural, optical and luminescence studies of CuO doped MgZn₂(PO₄)₂ nanopowder variants for solid state optical devices. *Next Materials* 2025;8:100819. <https://doi.org/10.1016/J.NXMATE.2025.100819>.
- [13] Oneto L, Navarin N, Micheli A, Pasa L, Gallicchio C, Bacciu D, et al. Informed machine learning for complex data. *Neurocomputing* 2026;669:132505. <https://doi.org/10.1016/J.NEUCOM.2025.132505>.
- [14] Li A, Xiao F, Xiao Z, Yan R, Li A, Lv Y, et al. Active learning concerning sampling cost for enhancing AI-enabled building energy system modeling. *Advances in Applied Energy* 2024;16:100189. <https://doi.org/10.1016/J.ADAPEN.2024.100189>.
- [15] Kumar JS, Lakshmi NR, Neeraja K, Devi TA, Swamy DT, Basha SA, et al. Spectroscopic and Optical Studies of Molybdenum Ions in CdO-SrO-B₂O₃-SiO₂ (CdSBSi) Glass System. *Asian Journal of Chemistry* 2025;37:1204–10. <https://doi.org/10.14233/AJCHEM.2025.33684>.
- [16] Cole S, Basha SA, Naveen Kumar B V, Srinivasarao K, Daniel K. Eu³⁺ Ion-doped Lithium Cadmium Orthophosphate Nanopowders for Reddish Orange-emitting Solid-state Lighting Applications. *Physical Chemistry Research* 2026;14:173–94. <https://doi.org/10.22036/PCR.2025.539879.2719>.
- [17] Namazi H, Pradhan AK, Frischer R, Socha L. Machine learning-guided electrolyte–interface co-design for Li-based anode-free batteries: Toward a predictive framework for stable metal deposition. *J Power Sources* 2026;690:240826. <https://doi.org/10.1016/J.JPOWSOUR.2026.240826>.
- [18] Khawaja Y, Shankar N, Qiqieh I, Alzubi J, Alzubi O, Nallakaruppan MK, et al. Battery management solutions for li-ion batteries based on artificial intelligence. *Ain Shams Engineering Journal* 2023;14:102213. <https://doi.org/10.1016/J.ASEJ.2023.102213>.
- [19] Balasubramanyam S V., Kumar BVN, Srinivasarao K, Basha SA, Cole S. Spectroscopic and photoluminescent studies of Dy³⁺ doped MgCdP₂O₇ nanopowders for white light emitting devices. *Journal of Materials Science: Materials in Electronics* 2026 37:14 2026;37:1098-. <https://doi.org/10.1007/S10854-026-17475-7>.
- [20] V.Balasubramanyam S, Srinivasarao K, Azad Basha Shaik, Sandhya Cole. Investigation on the structural and morphological properties of magnesium cadmium pyrophosphate nanopowders doped with Pr³⁺ ions. *Research Methodology in Materials Science* 2025;1:145–9.
- [21] Bakkar Siddique MA, Parves MS, Tarekuzzaman M, Kabir MR, Al-Saleem MSM, Al-Humaidi JY, et al. Exploring Optoelectronic Behavior and Solar Cell Efficiency of Double Halide Perovskites M₂KIrCl₆ (M = Cs, Rb) through DFT and SCAPS-1D. *Langmuir* 2025;41:19797–820. <https://doi.org/10.1021/ACS.LANGMUIR.5C01639>.
- [22] Hui Z, Wang M, Yin X, Wang Y, Yue Y. Machine learning for perovskite solar cell design. *Comput Mater Sci* 2023;226:112215. <https://doi.org/10.1016/J.COMMATSCI.2023.112215>.
- [23] Basha SA, Srinivasarao K, Daniel K, Naveen Kumar B V., Rajeev YN, Cole S. Orthorhombic crystal structured Dy³⁺ doped Li₂Cd₂(PO₄)₂ nanopowders for yellowish light emitting solid state lighting applications. *Chemical Papers* 2026 2026:1–18. <https://doi.org/10.1007/S11696-026-05178-6>.
- [24] Basha SA, Srinivasarao K, Daniel K, Kumar BVN, Cole S. Reddish Orange Emitting Sm³⁺ Ion on Lithium Cadmium Orthophosphate Nanopowders for Solid State Lighting Applications. *Asian Journal of Chemistry* 2026;38:2151–62. <https://doi.org/10.14233/AJCHEM.2026.36352>.

- [25] Li Z, Zhang Y, Zhou T, Jia G. Accelerating electrocatalyst design for CO₂ conversion through machine learning: Interpretable models and data-driven innovations 2024;1:100029. <https://doi.org/10.1016/J.YNEXS.2024.100029>.
- [26] Wang C, Wang B, Wang C, Li A, Chang Z, Wang R. A machine learning model with minimize feature parameters for multi-type hydrogen evolution catalyst prediction. Npj Computational Materials 2025 11:1 2025;11:111-. <https://doi.org/10.1038/s41524-025-01607-4>.
- [27] Rani GR, Srinivasarao K, Daniel K, Basha SA, Cole S. Exploring structural and optical characteristics of MnO₂doped MgZn₂(PO₄)₂ nanopowdersfor Optoelectronic devices. Journal of the Indian Chemical Society 2026:102957. <https://doi.org/10.1016/J.JICS.2026.102957>.
- [28] Srinivasarao Kuna, Shaik Azad Basha, Sandhya Cole. Methods for Designing and Improving Luminescent Nanophosphors for LED Applications Using Machine Learning. National Seminar on Transforming India; Innovations, Challenges and Opportunities n.d.;1:138–59.
- [29] Choudhary K, DeCost B, Chen C, Jain A, Tavazza F, Cohn R, et al. Recent advances and applications of deep learning methods in materials science. Npj Computational Materials 2022 8:1 2022;8:59-. <https://doi.org/10.1038/s41524-022-00734-6>.
- [30] Yasin Ahmed T, Aziz SB, M. A. Dannoun E. New photocatalytic materials based on alumina with reduced band gap: A DFT approach to study the band structure and optical properties. Heliyon 2024;10:e27029. <https://doi.org/10.1016/J.HELİYON.2024.E27029>.