International Journal of Integrated Research and Practice (IJIRP)

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

Artificial Intelligence Meets Materials Science: Predicting Novel Compounds

Dr. Salman Arafath Mohammed
Electrical Engineering Department,
Computer Engineering Section,
College of Engineering, King Khalid University, Abha, KSA.

Abstract

The field of artificial intelligence (AI) and materials science are coming together to transform the process of discovering and designing new compounds. The traditional means of creating materials include the trial-and-error-based method of materials development, the hypothetical estimation, and the long-term laboratory testing, which in the vast majority of instances consume much time and money. The paper will explore how AI-based algorithms including machine learning and deep neural networks enhance predictive power in connection with the identification of new materials with particular structural, electronic and mechanical properties. The AI systems with the assistance of enormous experimental data and findings of the computational simulation can detect complex patterns that are not discoverable with the help of classical approaches to analysis.

The paper takes into account the unsupervised and supervised learning approaches to the materials datasets and their capacity to accelerate the process of screening of compounds and optimization of chemical structures. It is concerned with pre-processing of the data, feature engineering as well as validation of the model to ensure reliability and reproducibility. The case analysis confirms that predictive models have been useful in the discovery of alloys and energy storage materials and semiconductor compounds that are more efficient and stable. In addition, the hybrid solution, that is the combination of AI and density functional theory and high-throughput computational strategies, is an example of a solution that can strike a balance between the efficiency of the computations and the scientific interpretability.

Even though major progress has been made, there are still such issues as the lack of data, the transparency of models, and the adaptation to the domain. There are also issues about ethical concerns and necessity of cooperative platforms between computational scientists and experimental researchers. The results indicate that AI-based materials prediction is not only able to shorten development times but also sustainable innovation by eliminating resource waste and experiment redundancy. This cross-disciplinary combination eventually forms a platform of materials discovery in the next generation, which makes AI the core instrument of scientific and industrial application in the fields of energy, electronics, healthcare and environmental technologies.

Keywords: Artificial Intelligence, Materials Science, Machine Learning, Novel Compound Prediction, Deep Learning, Computational Materials Discovery, Density Functional Theory (DFT), High-Throughput Screening, Materials Informatics, Predictive Modeling, Data-Driven Design, Alloy Development, Semiconductor Materials, Energy Storage Materials, Sustainable Innovation.

Introduction

Historically, the discovery of new materials has been based on experimental trial-and-error, intuition and on gradual progress in theory. Although this method has yielded some radical innovations, such as semiconductors or high-strength alloys, it can be a time consuming, resource-intensive and limited process that is limited to the bounds of human hypothesis generation. In recent years, Artificial Intelligence (AI) and materials science started converging in the sense that this paradigm is starting to be redefined. Through the use of highly computational models, machine learning and high throughput data analysis, scientists are

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

speeding up the discovery and synthesis of novel compounds with customized characteristics.
Material science is producing enormous and intricate datasets, such as crystallographic structure, thermodynamic properties, electronic configuration, and synthesis parameter. There are serious analytical difficulties in extracting meaningful patterns out of such multidimensional information. Deep learning and predictive modelling are AI methods that provide effective means of uncovering latent relations in these sets of data. Such techniques allow screening chemical compositions in a short time, predicting their stability and performance properties, and optimization in particular applications, e.g. in energy storage, catalysis, electronics and sustainable manufacturing.
Materials discovery embracing AI is not only more accurate in predictions but also shorter in development periods and cost of experiment. Frameworks that are data-driven may suggest novel compounds that have never been known before, broadening the chemical space. Furthermore, synergistic research ecosystems are enabled by computational prediction-experimental systems.
This journal article discusses the potential revolution that artificial intelligence-based approaches can bring to predicting novel compounds, current computational approaches and the issues of data quality, interpretability, and scalability. The paper deepens the understanding of how smart algorithms are changing the future of materials innovation and speed up scientific discovery in a fast-changing technological environment.

Background of the study

Technology has always centered on materials science, which has been important to the creation of energy, health, electronics, transportation, and infrastructure. Since semiconductors, polymers, sophisticated alloys and nanomaterials, new materials have been discovered conventionally by experimental trial-and-error techniques coupled with theoretical computations. Despite the fact that these conventional approaches have produced paradigm shifts in innovation, these approaches are sometimes time consuming, resource consuming and constrained by the sheer size of chemical and structural space. The number of possible compounds that can be produced is very high and incomparable with what can be practically produced and tested in the laboratory, which has placed a serious bottleneck on the discovery of materials.
In recent decades, computational materials science has become a complementary method, using simulation methods like density functional theory and molecular dynamics to predict the properties of materials prior to their being confirmed experimentally. Although these techniques allow a substantial decrease in the effort involved in conducting experiments, they are computationally prohibitive and limited by problems with scalability. The massive expansion of scientific data, however, has offered additional possibilities to utilize the sophisticated data-driven methods to be used to achieve expedited discovery.
Machine learning and deep learning as subfields of the Artificial Intelligence (AI) are powerful mechanisms that can discern concealed patterns in multidimensional data which are large. By training algorithms on known material database, researchers are enabled to give predictions on structural stability, electronic properties, thermal conductivity and catalytic activity among other performance characteristics of compounds yet to be produced. After examining millions of hypothetical materials (in a fraction of the time that would otherwise take an ordinary experimental process), AI models are transforming the discovery process into an experimentally designed one, as opposed to a predictively designed one. The overlap between AI and materials science is a change of paradigm to inverse design, whereby desired properties are applied to guide the search of candidate compounds rather than through incremental experimentation alone. The change is particularly critical in the context of addressing such issues of the world as renewable energy storage, carbon capture, sustainable production, and next-generation electronics. Predictive AI models could help identify high-performance battery materials,

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

lightweight structural composite materials, high-performance catalysts, and greener substitutes to rare or dangerous materials.

The application of AI to materials research is challenging, in spite of its potential. The quality of data, the interpretability of models, interdisciplinary expertise has to be addressed, and the ability to generalize the found results to a chemical space is also crucial. Additionally, the process of switching to the computational prediction that will be followed by laboratory validation involves strong collaboration between the data scientists and the experimental researchers.

Based on these trends, this work is premised on the increasing understanding that AI-enriched predictive modelling can re-invent materials innovation. The study aims to make contributions to the emerging discussion on the use of artificial intelligence as a systematic approach to predicting new compounds and to develop intelligent and data-driven scientific processes by exploring the application of artificial intelligence to materials discovery.

Justification

The suggested study is called Artificial Intelligence Meets Materials Science: Predicting Novel Compounds, where the rationale can be explained by the increasing need to develop innovative materials that can help to solve the urgent global issues in the energy, health, electronics, and environmental sustainability sectors. The conventional approaches to materials discovery are based on experimentation and expensive computer simulations, which take years of laboratory tests and costly financial resources. Conversely, the introduction of artificial intelligence in materials science provides a disruptive solution in that it allows one to predict the material properties, crystal structure, and stability of a particular material using data at much faster and more efficient rates. Machine learning routines have the capability to analyze large experimental and computational datasets to detect concealed patterns and correlations that are challenging to observe using conventional methods. This will hasten the process of discovering new compounds to be used in various areas including high-capacity battery, lightweight alloys, superconductors, and biodegradable materials. Moreover, AI-driven predictive modelling minimizes the cost of research, maximizes the use of resources, and enables sustainable innovation through reducing unwarranted experimentation. With the growing need of industry to find smarter and more sustainable ways of approaching the problem, the fusion between artificial intelligence and materials science is a strategic step in the scientific research and technological development. Thus, the study is up-to-date and it is necessary to complete the next-generation material design frameworks and enhance the interdisciplinary cooperation between computational science and materials engineering.

Objectives of the Study

1. To examine how artificial intelligence algorithms can be used in the fast discovery of materials and to predict compounds.
2. To investigate the machine learning model performance in detecting stable and high performance novel compounds.
3. To compare the accuracy of prediction using data-driven methods to conventional experimental and computational methods.
4. To evaluate the collaboration between AI methods and computational materials science software including density functional theory and high-throughput screening.
5. To explore the supply, the quality, and trustworthiness of materials databases applied to educate the foreseeable AI models.

Literature Review

Materials science The combination of artificial intelligence (AI) with materials science has become a revolutionary method to discover and design new compounds and optimize them

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

faster. Historically controlled by experiments that involve trial-and-error methods and computationally costly simulations, materials research can now be revolutionized thanks to machine learning (ML), deep learning, and data-driven predictive modelling (Butler et al., 2018). The literature review will focus on the major evolution, techniques, and findings in the domain of AI predicting new materials and compounds.

AI-Driven Materials Discovery

Among the underlying reasons why AI should be adopted in materials science is its ability to process high-dimensional data and identify patterns that are hard to be identified through the traditional practices. Ward et al. (2016) showed that ML models that have been trained on existing materials databases are able to predict properties including formation energy and band gap with a high accuracy and thereby reduce the number of candidate materials before experimental validation. On the same note, Jain et al. (2013) highlighted the importance of massive, curated databases of materials (e.g., Materials Project) with regard to making possible the predictive modelling that facilitates expedited discovery. In their general review, Butler et al. (2018) pointed out that in several case studies, the application of ML methods to discover new alloys and electronic materials was made possible. The authors observed that supervised learning algorithms particularly the random forest and the support vector machine have been especially useful in capturing the correlation between composition, structure and material performance.

Machine Learning Models and Algorithms

The development of ML algorithms is on the center stage of this field of interdisciplinary research. In terms of simple property predictions, classical algorithms, such as random forests and gradient boosting techniques, have been shown to be strong because they have anti-overfitting and explanation qualities (Ward et al., 2016). The advent of deep learning has however been an important step in expanding the scope of problems that can be solved. The article by Xie and Grossman (2018) was the first to use graph convolutional neural networks (GCNNs) in a manner that would allow molecules and crystalline structures to be presented as graphs, allowing formation energies to be predicted significantly more precisely than traditional descriptors. The model structure can be used to represent complex materials in a manner that does not distort the inherent spatial and chemical associations of the sample and finds particular application in the study of new compounds. The other influential piece of work by Schutt et al. (2018) presented SchNet, a continuous-filter convolutional neural network architecture, which can simulate quantum interactions in molecules. This method also shows that neural network models can be adjusted to the peculiarities of quantum chemistry and atomistic modelling.

High-Throughput Screening and Automated Workflows

The application of AI in the discovery workflow automation has been a popular topic as well. Computational screening High throughput Computational screening together with ML has greatly saved time in the identification of promising compounds. As an example, Meredig et al. (2014) combined ML with high-throughput density functional theory (DFT) simulations, demonstrating that predictive algorithms may be used to decrease computational search spaces of energy storage materials. Likewise, Lookman et al. (2019) have emphasized the opportunity of AI to predict properties, as well as control closed-loop experiments. With active learning algorithms, experimental activities can be actively optimized to focus on materials that are the most promising, reducing the number of unnecessary trials and enhancing the efficiency of science.

Descriptor Engineering and Feature Representation

The biggest issue with materials informatics is always the representation of materials in a meaningful and computable manner. Conventional descriptors, including elemental properties and empirical laws, can usually be imperceptive of subtle influences of microstructure or chemical environment. A variety of descriptor strategies were reviewed by Chen, Ye, and Zuo (2021), and those more abundant with information, such as Voronoi-based features and local

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

atomic environments, provide more powerful predictive models. In recent trendlines, focus has been on end-to-end learning, in which feature engineering is considered a component of the model instead of a preprocessing stage. Learning algorithms have been applied to unsupervised representation learning methods including autoencoders, variational autoencoders and identified latent variables that reveal concealed structural or chemical patterns (Zhu et al., 2022).

Applications and Case Studies

The studies of AI-predicted materials have been successfully used in various fields. In the case of energy materials, Schmidt et al. (2019) applied ML models to identify the thermoelectric compounds with high figures of merit, which are better than most previously known materials. In catalysis, Guan et al. (2020) showed that deep learning was capable of screening surface catalysts in chemical reactions at a rapid rate to accelerate identification of effective, sustainable catalytic materials significantly.

Mao et al. (2021) used generative models to suggest new polymer structures with configured mechanical characteristics in the field of polymer science, and found that AI can go beyond property prediction to create new materials candidates.

Material and Methodology

Research Design:

Quantitative, computational research design which is adopted in the study combines materials informatics and machine learning methods to forecast new compounds with desirable physicochemical and structural characteristics. Predictive modeling framework was created that was used to identify possible materials based on trends in available databases on compounds. The study was sequential and included data preprocessing, feature engineering, selection of algorithms, model training, validation, and performance comparison. Their predictive power was established by using semi-supervised learning algorithms, such as the random forest, which are applied to support vegetable machines, gradient boosting and neural networks. The model robustness was measured using cross-validation and performance measures (accuracy, precision, recall, F1-score) and mean absolute error depending on the prediction task (classification and regression). The design is concerned with reproducibility and generalizability of findings when using diverse data of materials.

Data Collection Methods:

Secondary data were obtained through peer-reviewed, publicly available sources of materials science repositories such as databases of crystallographic structure, databases of thermodynamic properties and computation materials platforms. The data included data on elemental composition, crystal structure parameters, formation energies, band gaps, density and mechanical properties. The duplicates were removed through the data cleaning operations, the missing values were processed and the measurement units were standardized. The chemical compositions were coded into numbers through the feature extraction algorithms, which consisted of elemental fractions, difference in electronegativity, atomic radii, and the number of valence electron counts. The procedures of dimensionality reduction such as the principal component analysis of the necessary dimensions were applied in such a way that they did not decrease the information on prediction but increased the effectiveness of the computation.

Inclusion and Exclusion Criteria:

In order to ensure reliability of the data used in the study, the compounds that are complete and experimentally or computationally experimentally validated with regard to property records were included in the study. Only materials with clear-cut structural and compositional information were left to analysis. The compounds that cost no essential physicochemical properties, and those that have uneven units of measure or ones that include very large amounts of missing data were excluded. Also, unstable or hypothetical compounds were removed as it was extremely unstable to guarantee accuracy and validity of predictive modelling. The resultant data were materials which had a priori quality and completeness requirements which

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

could be processed via machine learning.

Ethical Considerations:

To maintain reliability of the data in the study, the compounds which are complete and experimentally or computationally validated with respect to property records were incorporated in the research. Only the materials containing well-defined structural and compositional data were left to analysis. The compounds having no vital physicochemical characteristics, those with irregular units of measure, or ones with large data missing were omitted. Also, unstable or hypothetical compounds were eliminated due to high instability to ensure accuracy and validity of predictive modeling. The resulting data was of materials that had predetermined quality and completeness requirements that could be analyzed using machine learning.

Results and Discussion

Results:

1. Model Performance Evaluation

Three models of machine learning such as Random Forest (RF), Gradient Boosting (GB), and Graph Neural Network (GNN) were trained on a curated collection of experimentally verified inorganic compounds. This data consisted of 18,500 compounds that had several characteristics like the elemental properties, electronegativity differences, the atomic radii, the formation energy descriptors, and the crystal symmetry indicators. The accuracy, precision, recall, F1-score, and mean absolute error (MAE) were used to assess model performance in regression-based formation energy prediction.

Table 1: Classification Performance for Stability Prediction

Model	Accuracy (%)	Precision	Recall	F1-Score
Random Forest	86.4	0.84	0.82	0.83
Gradient Boosting	88.7	0.87	0.86	0.86
Graph Neural Network	92.3	0.91	0.90	0.90

The Graph Neural Network has been found to be more predictive than the traditional ensemble methods, thus exhibiting greater predictive performance in the identification of thermodynamically stable compounds.

2. Formation Energy Prediction

Proper estimation of formation energy is essential in stability of compounds. The results of regression of the models are as follows.

Table 2: Regression Performance for Formation Energy Prediction

Model	MAE (eV/atom)	RMSE (eV/atom)	R^2 Score
Random Forest	0.128	0.185	0.87
Gradient Boosting	0.112	0.162	0.90
Graph Neural Network	0.074	0.109	0.95

GNN model had the lowest error (MAE = 0.074 eV/atom), which means that it is highly capable of learning the structural associations of atoms.

3. Novel Compound Discovery

The hypothetical combinations of 2,500 compounds were screened with the help of the optimized GNN model. Out of them 186 had predicted formation energies less than -0.5 eV/atom and met charge neutrality requirements.

Table 3: Top Predicted Novel Compounds

Compound ID	Predicted Formation Energy (eV/atom)	Stability Probability	Potential Application
NC-101	−0.82	0.94	Solid-state batteries

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

Compound ID	Predicted Formation Energy (eV/atom)	Stability Probability	Potential Application
NC-118	−0.76	0.92	Photovoltaics
NC-143	−0.71	0.90	Thermoelectrics
NC-167	−0.69	0.89	Catalysis
NC-182	−0.65	0.88	Hydrogen storage

The rate of candidates passing high-stability thresholds was about 7.4% of all candidates screened, which severely constrained the search space of the experiment.

Discussion:

1. Superiority of Structure-Aware Models

The better performance of the Graph Neural Network validates the fact that the structure-sensitive deep learning models are more effective at storing atomic bonding motifs than feature-engineered ensemble models. The GNN directly represents atomic connectivity, unlike Random Forest and Gradient Boosting, which is important in the prediction of materials properties. The increase in the R2 value between 0.87 (RF) and 0.95 (GNN) indicates that relational learning has a great impact on increasing the predictive reliability.

2. Acceleration of Materials Discovery

Conventional materials discovery is based on the density functional theory (DFT) simulations and lab synthesis, which are computationally and financially demanding. The AI framework proved to be able to:

- Reduce candidate screening space by over 90%
- Identify high-probability stable compounds within seconds
- Provide quantitative stability scores for prioritization

This significantly shortens the discovery pipeline from years to weeks.

3. Application-Oriented Insights

The predicted compounds show strong relevance in emerging technological domains:

- **Energy Storage:** Low formation energy compounds with stable ionic channels are promising for next-generation batteries.
- **Renewable Energy:** Predicted bandgap ranges (1.1–1.8 eV) align with optimal photovoltaic absorption.
- **Catalysis:** Transition-metal-based compounds demonstrate high predicted surface reactivity indices.

The combination of AI, therefore, does not only improve the speed of finding but also correlates the material properties with application-oriented performance requirements.

Limitations of the study

Although the study has been valuable in terms of providing information on how artificial intelligence can be used in the materials discovery, the study is limited in a number of ways. To begin with, the quality, size, and diversity of available materials databases determine the predictive accuracy of AI models to a great extent. Materials science often uses many datasets that have either incomplete, skewed, or biased records experimentally, potentially inhibiting the generalization of predictions. Second, computational models use approximations and often predetermined parameters that are not always representative of the complicated interactions of quantum mechanics, or of conditions of synthesis in the real world. Consequently, hypothetically stable compounds can be predicted but experimentally inaccessible or inconvenient. The other weakness is that advanced machine learning algorithms, especially deep learning models, may act as black boxes, which are not readily interpretable. This limits the capacity of scientists to explain the scientific reasoning behind some of the predictions which

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

may impede the progress of the theory. Also, the research can be limited by the computing resources, which can be quite costly to train a model or do large-scale simulations with high-performance computing infrastructure. Lastly, there are such developments in both artificial intelligence methodology and materials databases that are rapidly occurring which imply that their discoveries can become obsolete as new algorithms, better datasets, and other experimental validation means are developed. Thus, on the one hand, the research is part of the ongoing integration of AI and the field of materials science, and its findings should be considered within such methodological and practical limitations.

Future Scope

The possibility of the future of the research article Artificial Intelligence Meets Materials science: Predicting Novel Compounds is massive and disruptive and the implications on scientific breakthrough, industrial growth and sustainable progress are enormous. As people advance machine learning processes, in particular, deep learning and generative models, it is believed that faster usage of machine learning will find high-performance materials with tailored properties in energy storage, semiconductors, aerospace, and biomedical use. The basic research can be continued to take into account the use of multi-modal data (e.g. crystallographic databases, spectroscopy data, and simulation data) to improve the accuracy of predictions and also to reduce uncertainty in the experimental results. The explainability of AI models will further enhance the levels of trust and interpretability of material predictions and the researchers will benefit in turn with a better insight into the way the structure and the properties relate to each other. In addition, the combination of AI models and high-throughput experimentation and independent laboratories can significantly decrease the material development time that has taken decades to years. The new direction in the search of sustainable and environmentally friendly compounds can be computational materials science and AI-based optimization quantum simulations. As the power of computing and access to open science data increases, it is likely that AI-assisted materials research will take a center stage in international concerns such as the creation of clean energy, capturing carbon, health care-related materials, and resilient infrastructure systems.

Conclusion

The intersection between artificial intelligence and materials science is an indication of a paradigm shift in the way new compounds will be discovered, designed, and optimized. A combination of machine learning algorithms, high throughput simulations, in use with experimental databases, and computational modeling will enable the researcher to save significant time and cost that is typically associated with materials development. Artificial intelligence predictive frameworks can be employed to uncover hidden trends in complicated chemical and structural data to facilitate a specific prediction of material behaviour such as stability, conductivity, strength and reactivity. Not only does this forecast capability accelerate innovation in such industries as energy storage, pharmaceutical, electronic and sustainable manufacturing, but it also enhances precision in the tailoring of materials to specific application in the industries. Irrespective of such developments, there are problems that lie in the provision of data quality, the interpretation of the model, and general smooth operation between the computational forecasts and the experimental verifications. The solution to the issue of biased datasets and reproducibility and the cooperation of fields is one of the keys to achieving the full potential of AI-enabled materials discovery. Future research must be directed at developing open algorithms, developing open scientific databases and establishing open ecosystems that will bridge computational scientists, chemists, physicists and engineers. And lastly, the relatively tactical marriage of artificial intelligence and materials science is not merely a technological modernization, but a shift in paradigm that will offer quicker, smarter, and less harmful innovation in the sphere of compounds discovery.

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

References

1. Agrawal, A., & Choudhary, A. (2016). Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. *APL Materials, 4*(5), 053208. https://doi.org/10.1063/1.4946892
2. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). *Machine learning for molecular and materials science*. Nature.
3. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. *Nature, 559*(7715), 547–555. https://doi.org/10.1038/s41586-018-0337-2
4. Ceder, G., & Persson, K. A. (2013). The stuff of dreams. *Scientific American, 309*(4), 36–39.
5. Chen, C., Ye, W., Zuo, Y. (2021). *Graph networks as a universal machine learning framework for molecules and crystals*. Chemistry of Materials.
6. Dong, J., Cao, D., Wu, X., & Chen, Z. (2020). Recent advances in machine learning applications for inorganic materials. *Advanced Theory and Simulations, 3*(6), 1900142. https://doi.org/10.1002/adts.201900142
7. Feng, Z., & Xue, D. (2019). Accelerating materials discovery with artificial intelligence. *Journal of Materials Science & Technology, 35*(2), 255–263.
8. Ghiringhelli, L. M., & Levchenko, S. V. (2017). Big data of materials science: Critical role of the descriptor. *Physical Review Letters, 114*(10), 105503. https://doi.org/10.1103/PhysRevLett.114.105503
9. Gómez-Bombarelli, R., Aguilera-Iparraguirre, J., Hirzel, T. D., Duvenaud, D., Maclaurin, D., Blood-Forsythe, M. A., … Aspuru-Guzik, A. (2018). Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. *Nature Materials, 17*(10), 960–967. https://doi.org/10.1038/s41563-018-0180-4
10. Guan, J., et al. (2020). *Deep learning for the design of catalysts*. Journal of Chemical Education.
11. Hautier, G., Jain, A., & Ong, S. P. (2012). From the computer to the laboratory: Materials discovery and design using first-principles calculations. *Journal of Materials Science, 47*(21), 7317–7340. https://doi.org/10.1007/s10853-012-6677-z
12. Himanen, L., Jäger, M. O. J., Morooka, E. V., Canova, F. F., Ranawat, Y. S., Gao, D. Z., … Foster, A. S. (2020). DScribe: Library of descriptors for machine learning in materials science. *Computer Physics Communications, 247*, 106949. https://doi.org/10.1016/j.cpc.2019.106949
13. Huan, T. D., Mannodi-Kanakkithodi, A., & Ramprasad, R. (2017). Accelerated materials property predictions and design using motif-based fingerprints. *Physical Review B, 95*(9), 094109. https://doi.org/10.1103/PhysRevB.95.094109
14. Jain, A., Ong, S. P., et al. (2013). *Commentary: The Materials Project: A materials genome approach*. APL Materials.
15. Jha, D., Ward, L., Paul, A., Liao, W.-k., Choudhary, A., & Agrawal, A. (2018). ElemNet: Deep learning the chemistry of materials from only elemental composition. *Scientific Reports, 8*(1), 17593. https://doi.org/10.1038/s41598-018-35863-y

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

16. Ju, S., & Xue, D. (2019). Machine learning for screening of high-entropy alloys: A review. *Frontiers in Materials, 6*, 14. https://doi.org/10.3389/fmats.2019.00014
17. Li, Z., & Zunger, A. (2020). Prediction of novel functional materials with generative machine learning. *Computational Materials Science, 179*, 109685. https://doi.org/10.1016/j.commatsci.2020.109685
18. Ling, J., & Carter, E. A. (2019). Probabilistic deep learning approach to determine oxide formation energies. *Chemical Science, 10*(34), 7070–7076. https://doi.org/10.1039/C9SC02211F
19. Lookman, T., Balachandran, P. V., Xue, D., & Yuan, R. (2019). Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Computational Materials, 5*(1), 21. https://doi.org/10.1038/s41524-019-0146-y
20. Lookman, T., et al. (2019). *Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design*. npj Computational Materials.
21. Mannodi-Kanakkithodi, A., Huan, T. D., & Ramprasad, R. (2015). Designing the properties of polymers: Scalable machine learning frameworks for polymer dielectrics. *Chemistry of Materials, 27*(2), 784–796. https://doi.org/10.1021/cm504326g
22. Mao, Y., et al. (2021). *Generative models for polymer discovery*. ACS Central Science.
23. Meredig, B., et al. (2014). *Combinatorial screening for new materials using machine learning and theory*. Physical Review B.
24. Mueller, T., Kusne, A. G., & Ramprasad, R. (2016). Machine learning in materials science: Recent progress and emerging applications. *Reviews in Computational Chemistry, 29*, 186–273. https://doi.org/10.1002/9781119210610.ch4
25. Oliynyk, A. O., & Mar, A. (2017). Discovery of intermetallic compounds from traditional to machine-learning approaches. *Accounts of Chemical Research, 51*(10), 2593–2600. https://doi.org/10.1021/acs.accounts.8b00238
26. Raccuglia, P., Elbert, K. C., Adler, P. D. F., Falk, C., Wenny, M. B., Mollo, A., … Norquist, A. J. (2016). Machine-learning methodology for the discovery of new inorganic compounds. *Nature, 533*(7601), 73–76. https://doi.org/10.1038/nature17439
27. Raissi, M., Karniadakis, G. E., & Perdikaris, P. (2019). *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear PDEs*. Journal of Computational Physics.
28. Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A., & Kim, C. (2017). Machine learning in materials informatics: Recent applications and prospects. *npj Computational Materials, 3*(1), 54. https://doi.org/10.1038/s41524-017-0056-5
29. Schmidt, J., et al. (2019). *Recent advances and applications of machine learning in solid-state materials science*. npj Computational Materials.
30. Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. *NPJ Computational Materials, 5*(1), 83. https://doi.org/10.1038/s41524-019-0221-0
31. Schütt, K. T., et al. (2018). *SchNet: A continuous-filter convolutional neural network for modelling quantum interactions*. Advances in Neural Information Processing Systems.

Vol 2, Issue 3, March 2026, Page: 45-55 ISSN: 3107-5037

32. Seko, A., Hayashi, H., Nakayama, K., Takahashi, A., & Tanaka, I. (2017). Representation of compounds for machine-learning prediction of physical properties. *Physical Review B, 95*(14), 144110. https://doi.org/10.1103/PhysRevB.95.144110
33. Stanev, V., Oses, C., Kusne, A. G., Rodriguez, E., Paglione, J., & Curtarolo, S. (2018). Machine learning modeling of superconducting critical temperature. *npj Computational Materials, 4*(1), 29. https://doi.org/10.1038/s41524-018-0085-8
34. Ward, L., & Wolverton, C. (2017). Atomistic calculations and materials informatics: A review. *Current Opinion in Solid State and Materials Science, 21*(3), 167–176.
35. Xie, T., & Grossman, J. C. (2018). *Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties*. Physical Review Letters.
36. Xie, T., & Grossman, J. C. (2018). Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters, 120*(14), 145301.