Journal of Intelligent Pervasive and Soft Computing

journal homepage: https://scirep.institute/journals/index.php/jipsc/

Artificial Intelligence-based Modeling Mechanisms for Material Analysis and Discovery

Imran [1,2] and Do Hyeun Kim [2,*]

[1] Engineering Department, Jeju National University, Jeju-si 63243, Korea.
[2] Department of Research and Innovation, SCIREP Institute of Scientific Research and Entrepreneurship, Islamabad 46000, Pakistan.
[*] Authors to whom correspondence should be addressed.

Abstract

Artificial intelligence-based materials application is one of the hot topics in the field of materials science. Materials are widely used in the space industry, cutting tools, thermal and electrical insulators, and refractory materials. The conventional experiments and statistical approaches usually require more resources and time. Thus, the need for Artificial intelligence applications in the simulation and exploration of novel materials is increasing. Recently, AI has been applied to materials for improving efficiency and prediction accuracy; however, there are many limitations due to the lack of benchmark datasets, advanced pre-processing mechanisms, prediction modelling mechanisms, and simulation tools in the materials literature. Thus it is challenging to identify optimal learning models, including algorithm selection, the architecture of models, data processing, and simulation mechanisms. In this paper, we attempt to review experimental and computational data-based AI mechanisms. Furthermore, the current research status is analyzed for using artificial intelligence techniques in material simulation tools and discovering new materials. Finally, we present research issues of AI-based application realization in materials science.

Keywords*: Material analysis; Modeling mechanisms; AI simulations; Material analysis;*

1. Introduction

Artificial intelligence (AI) applications are bringing drastic changes in every field of life, especially in the technological areas [1-5]. Due to its successful adaptability for high efficiency and performance, the need for AI in the simulation and exploration of novel ceramic materials is increasing. Materials design based on AI analysis is foreseen to lead to innovative materials and reduce the development cost in terms of time and resources [6-8]. However, the scientific community identified many limitations of advanced materials discovery and application based on AI and advanced machine learning techniques. For instance, there are several issues related to computational simulation, and the materials' structures require high-performance index properties. Therefore, advanced materials research based on the convergence between AI techniques and experimental mechanisms is needed to produce the basic comprehension of the input parameters conditions and performance index properties [9]. Fig 1 presents the general AI Mechanism in ceramic materials science. In the first step, raw data is prepared for model building using pre-processing and feature engineering techniques. The second step is building an AI model using AI-based learning

algorithms. Finally, the performance evaluation of the model and interpretation of the model knowledge into input parameters and performance index properties for assessment of the materials.

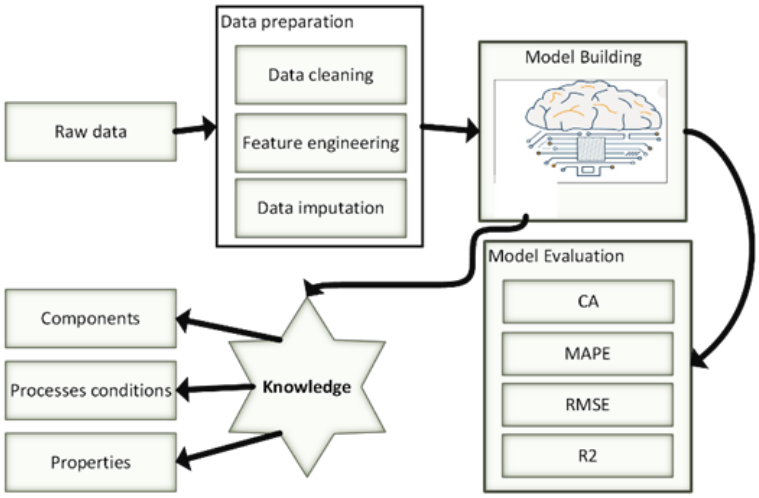

Fig. 1. The general AI Mechanism in materials science for modelling and simulation

2. Modelling Mechanisms

The research interest in modelling mechanisms has increased over the past few decades for interdisciplinary research and applicatoins[10]. Due to interdisciplinary research interest, material applications based on AI mechanisms have emerged recently with increased experimental and computational data. One important task of material science-based AI applications is to use existing material data to predict the properties of new materials by using data science methods and mathematics [11]. The first step is to build a descriptor model that can predict the property of interest based on a known set of input material-specific features. For instance, one of the vital descriptor models where input variables are material structure features is called quantitative structure-property relationship (QSPR). In the input and output of material properties, a complex relationship is challenging to handle by traditional linear and non-linear correlation methods. However, these complex relationships can now be efficiently modelled by ML methods[12].

After performing descriptor modelling analysis and familiarizing with data, a model is applied to predict material behaviour using material analysis models based on AI or statistics. Machine learning(ML) models are intelligent and improve performance automatically by experience using training data to find patterns from data. Below is different ML algorithms categorized according to their type.Table1 presents the Predictive mechanism used in material science applications.

Table 1Predictive mechanism used in material science applications.

Methods	Category
Least-squares regression	Regression
Kernel ridge regression	Regression
Kriging or Gaussian process regression	Regression
Artificial Neural Network	Regression Classification
Support Vector Machine	Regression Classification
Decision tree	Classification
Random forest	Classification
k-nearest neighbors	Classification

Naive Bayes	Classification

For instance, Naive BayesClassification ML applications in material research are the property prediction of steel strength [13], alloys [14], and bandgaps[15]. Applications in the identification of stable materials are catalytic activities [16] and acid dissociation constants [17], [18]. Polymer-based material research applications are polymer dielectrics [19] and mixed oxide catalysts [20].

Supervised learning finds a function that maps input data to output property, while unsupervised learning finds the relationship among the data. Clustering in unsupervised learning divides a dataset into different categories such that the data points in the same group or cluster are more similar than those in other clusters. Clustering is best to extract physical insights from data and find new promising materials based on comparative studies. K-means, Hierarchical clustering, and hidden Markov modelling are popular clustering algorithms[21].

The scientific community also reviews ML tools for material discovery and modern applications of various ML strategies. However, a few published reviews on AI for materials have been published, emphasizing a single material system or individual methods. Therefore, the research focus on an application-based perspective in AI-enhanced material discovery. AI techniques reviewed in the paper include property prediction, characterization, synthesis, and theory paradigm discovery.

3. Simulation Tools

There have been many machine learning-based simulation tools developed in the literature. In this section, we briefly explore simulation tools published in highly reputed journals such as Nature. A deep understanding of material chemistry from only its elemental composition is compared with conventional ML approaches for predicting material properties from elemental compositions[22]. ElemNet is used to automatically capture the physical and chemical interactions and similarities through the design and implementation of deep neural networks. ElemNet enables rapid and robust screening in a large combinatorial space for new material candidates. Comparison of deep learning approach with conventional ML approach for prediction of materials properties. There are hundreds of thousands of chemical systems predicted by ElemNet that may contain yet-undiscovered compounds. By using deep learning, such manual feature engineering requiring domain knowledge can be avoided, resulting in better results, even with only a few thousand training samples.

An open-source software platform called matminer was developed to aid in data-driven analyses and predictions of materials properties. Matminer offers python modules for importing large data sets from external data sources. These data sources are based on Citrination, Materials Project, Materials Platform for Data Science, and Materials Data Facility databases. Moreover, it provides an API for implementing code from an extensive feature extraction library developed by the materials science community[23]. Table 2 lists simulation tools for materials analysis of material structures and properties publicly.

Table 2 Publicly accessible simulation tools for materials analysis

Name	Description
AFLOW	Online applications for property predictions using machine learning
CALPHAD	Computer Coupling of Phase Diagrams and Thermochemistry
Matminer	Data source, descriptive and predictive analysis
ElemNet	Deep learning-based mechanism
ChemSpider	Search engine for Chemistry's structure database
Citrination	AI-Powered Materials Data Platform

4. Conclusion

AI-based mechanisms in materials science have received notable concentration from material science researchers and interdisciplinary research communities. AI and its subfields, such as machine learning, have been proven excellent techniques for analyzing big data retrieved from material databases and datasets. These tools provide correlations between many complexes and interrelated structures of materials composition. This review paper summarizes material science applications based on AI-based modelling techniques and materials simulation tools. AI mechanisms such as deep learning show significant improvements and potential for the prediction of compositions, processes conditions, and performance properties of materials to reveal the changes in specific parameters for modelling materials' behaviour. Moreover, the discussion has been presented on AI techniques is preset, showing the significance of advanced AI mechanism-based models in designing and optimizing properties prediction for advanced materials discovery. For future work, we will develop material authoring and analysis framework based on composition assessment and machine learning for efficient material discovery.

5. References

1. Iqbal, N., & Kim, D. H. (2022). Iot task management mechanism based on predictive optimization for efficient energy consumption in smart residential buildings. Energy and Buildings, 257, 111762.
2. Jamil, F., & Kim, D. (2021). An ensemble of prediction and learning mechanism for improving accuracy of anomaly detection in network intrusion environments. Sustainability, 13(18), 10057.
3. Imran, Imran, et al. "Using Machine Learning Algorithms for Housing Price Prediction: The Case of Islamabad Housing Data." Soft Computing and Machine Intelligence 1.1 (2021): 11-23.
4. Imran, M., Zaman, U., Imtiaz, J., Fayaz, M., & Gwak, J. (2021). Comprehensive survey of iot, machine learning, and blockchain for health care applications: A topical assessment for pandemic preparedness, challenges, and solutions. Electronics, 10(20), 2501.
5. Zaman, U., Mehmood, F., Iqbal, N., Kim, J., & Ibrahim, M. (2022). Towards Secure and Intelligent Internet of Health Things: A Survey of Enabling Technologies and Applications. Electronics, 11(12), 1893.
6. Qayyum, F., Kim, D. H., Bong, S. J., Chi, S. Y., & Choi, Y. H. (2022). A Survey of Datasets, Pre-processing, Modeling Mechanisms, and Simulation Tools Based on AI for Material Analysis and Discovery. Materials, 15(4), 1428.
7. Olson, G.B., 2000. Designing a new material world. Science, 288(5468), pp.993-998.
8. Sitek, W., Trzaska, J. and Dobrzański, L.A., 2006. An artificial intelligence approach in designing new materials. Journal of Achievements in Materials and Manufacturing Engineering, 17(1-2).
9. Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A. and Kim, C., 2017. Machine learning in materials informatics: recent applications and prospects. npj Computational Materials, 3(1), pp.1-13.
10. Iqbal, N., Khan, A. N., Rizwan, A., Qayyum, F., Malik, S., Ahmad, R., & Kim, D. H. (2022). Enhanced time-constraint aware tasks scheduling mechanism based on predictive optimization for efficient load balancing in smart manufacturing. Journal of Manufacturing Systems, 64, 19-39.
11. Virshup AM, Contreras-García J, Wipf P, Yang W, Beratan DN. Stochastic voyages into uncharted chemical space produce a representative library of all possible drug-like compounds. J Am Chem Soc 2013;135(19):7296–303. link1
12. T. Zhou, Z. Song, and K. Sundmacher, 'Big Data Creates New Opportunities for Materials Research: A Review on Methods and Applications of ML for Materials Design', Engineering, vol. 5, no. 6, pp. 1017–1026, Dec. 2019, doi: 10.1016/j.eng.2019.02.011.
13. Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation. APL Mater 2013;1(1):011002. link1
14. Michalski RS, Carbonell JG, Mitchell TM, editors. ML: an AI approach. Berlin: Springer-Verlag; 2013. link1
15. Agrawal A, Deshpande PD, Cecen A, Basavarsu GP, Choudhary AN, Kalidindi SR. Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters. Integr Mater Manuf Innovation 2014;3:8. link1

16. Karak SK, Chatterjee S, Bandopadhyay S. Mathematical modeling of the physical and mechanical properties of nano-Y2O3 dispersed ferritic alloys using evolutionary algorithm-based neural network. Powder Technol 2015;274:217–26. link1
17. Pilania G, Mannodi-Kanakkithodi A, Uberuaga BP, Ramprasad R, Gubernatis JE, Lookman T. ML bandgaps of double perovskites. Sci Rep 2016;6:19375. link1
18. Jinnouchi R, Asahi R. Predicting catalytic activity of nanoparticles by a DFTaided machine-learning algorithm. J Phys Chem Lett 2017;8(17):4279–83. link1
19. Zhou T, Jhamb S, Liang X, Sundmacher K, Gani R. Prediction of acid dissociation constants of organic compounds using group contribution methods. Chem Eng Sci 2018;183:95–105. link1
20. Aghaji MZ, Fernandez M, Boyd PG, Daff TD, Woo TK. Quantitative structure– property relationship models for recognizing metal organic frameworks (MOFs) with high CO2 working capacity and CO2/CH4 selectivity for methane purification. Eur J Inorg Chem 2016;2016(27):4505–11. link1
21. Sharma V, Wang C, Lorenzini RG, Ma R, Zhu Q, Sinkovits DW, et al. Rational design of all organic polymer dielectrics. Nat Commun 2014;5:4845. link1
22. Jha, D., Ward, L., Paul, A., Liao, W.K., Choudhary, A., Wolverton, C. and Agrawal, A., 2018. Elemnet: Deep learning the chemistry of materials from only elemental composition. Scientific reports, 8(1), pp.1-13.
23. Ward, L., Dunn, A., Faghaninia, A., Zimmermann, N. E., Bajaj, S., Wang, Q., ... & Jain, A. (2018). Matminer: An open-source toolkit for materials data mining. Computational Materials Science, 152, 60-69.