Global Journal of Research in
Multidisciplinary Studies
Journal homepage: https://gsjournals.com/gjrms/
ISSN: 2980-4191 (Online)

(REVIEW ARTICLE)

Check for updates

Catalysis 4.0: A framework for integrating machine learning and material science in catalyst development

Matluck Ayomide Afolabi [1,*], Henry Chukwuemeka, Olisakwe [2] and Thompson Odion Igunma [3]

[1] *The University of Tennessee, Knoxville, USA.*
[2] *Department of Mechanical Engineering, Nnamdi Azikiwe University, Awka Nigeria.*
[3] *Independent Researcher, Florida, USA.*

Global Journal of Research in Multidisciplinary Studies, 2024, 02(02), 038–046

Publication history: Received on 19 October 2024; revised on 20 November 2024; accepted on 23 November 2024

Article DOI: https://doi.org/10.58175/gjrms.2024.2.2.0053

Abstract

This paper introduces Catalysis 4.0, a comprehensive framework for leveraging artificial intelligence (AI) and machine learning (ML) in the development of catalytic materials. The framework is structured around three key components: data-driven material discovery, virtual testing environments, and adaptive feedback loops. Data-driven material discovery utilizes AI algorithms to predict catalytic performance based on extensive material properties and high-quality datasets. Virtual testing environments provide simulation platforms to evaluate catalyst efficiency under various industrial conditions. This significantly reduces traditional experimentation time and costs. Adaptive feedback loops integrate real-time industrial data, enabling continuous refinement and improvement of AI models. This approach accelerates the catalyst development cycle, enhancing efficiency and reliability. The paper discusses the potential impact of this framework on industries such as refining and pharmaceutical manufacturing, where AI-driven catalyst design can enhance performance and reduce costs. Future directions for research include improving data collection techniques, developing sophisticated AI algorithms, and fostering interdisciplinary collaboration. The framework's implementation can revolutionize catalyst development, driving innovation and sustainability in critical industrial sectors.

Keywords: Catalysis 4.0; Artificial Intelligence; Machine Learning; Catalyst Development; Virtual Testing; Adaptive Feedback Loops

1. Introduction

1.1. Overview of Catalyst Development

Catalysts are substances that increase the rate of chemical reactions without being consumed in the process. They play a critical role in various industrial applications, significantly impacting sectors such as petrochemicals, pharmaceuticals, and environmental management (Hughes, Haque, Northey, & Giddey, 2021). The history of catalyst development dates back to the early 19th century when chemists like Humphry Davy and Jöns Jacob Berzelius first observed certain substances' acceleration of chemical reactions (Johnson, 2022). The advent of catalytic processes revolutionized industries by enabling more efficient and selective chemical reactions, leading to cost savings and reduced environmental impact (Rubtsov, Troshin, & Alymov, 2023).

Over the years, catalysts have evolved from simple natural minerals to highly sophisticated engineered materials. Traditional catalyst development involves a trial-and-error approach, which is time-consuming and resource-intensive. Scientists synthesize and test a multitude of materials to find optimal catalysts, which can take years of research and

* Corresponding author: Matluck Ayomide Afolabi

significant financial investment (Suwardi et al., 2022). Despite these challenges, catalysts remain indispensable due to their ability to enhance reaction rates, improve product yields, and lower energy consumption, making processes more sustainable and economically viable.

1.2. Emergence of Catalysis 4.0

The industrial landscape is undergoing a transformation with the advent of the Fourth Industrial Revolution, characterized by the fusion of digital, biological, and physical innovations (Sahai & Rath, 2021). In this context, "Catalysis 4.0" refers to the integration of advanced technologies such as artificial intelligence (AI) and machine learning (ML) in catalyst development. This new paradigm aims to overcome the limitations of traditional methods by leveraging data-driven approaches to accelerate the discovery and optimization of catalytic materials (Rane, Kaya, & Rane, 2024).

AI and ML algorithms have the potential to revolutionize catalyst design by analyzing vast datasets to identify patterns and correlations that human researchers might overlook. These technologies can predict the performance of new catalytic materials based on their properties, significantly reducing the time and cost associated with experimental testing (Benavides-Hernández & Dumeignil, 2024). Additionally, virtual testing environments allow for the simulation of catalyst behavior under various industrial conditions, providing valuable insights into their efficiency and stability without the need for extensive physical trials (Al-Akayleh, Ali Agha, Abdel Rahem, & Al-Remawi, 2024).

The increasing availability of computational power and advanced data analytics tools drives the emergence of Catalysis 4.0. High-throughput experimentation, combined with AI and ML, enables researchers to explore a broader range of materials and reaction conditions than ever before. This approach enhances the efficiency of catalyst development and opens up new possibilities for designing catalysts with tailored properties for specific applications, ultimately driving innovation in critical industries (Papadimitriou, Gialampoukidis, Vrochidis, & Kompatsiaris, 2024).

1.3. Objective of the Paper

This paper introduces a conceptual framework for leveraging AI and ML to accelerate the discovery and optimization of catalytic materials. The proposed framework consists of several key components, each addressing a specific aspect of the catalyst development process. By integrating data-driven material discovery, virtual testing environments, and feedback loops for continuous improvement, this framework aims to provide a comprehensive approach to Catalysis 4.0.

The first component of the framework focuses on data-driven material discovery, utilizing AI algorithms to predict catalytic performance based on material properties. This approach involves the collection and management of high-quality data, as well as the development of predictive models that can identify promising candidates for further testing.

The second component is the virtual testing environment, a proposed simulation platform designed to evaluate catalyst efficiency under various industrial conditions. By simulating different reaction environments, this platform allows researchers to assess the performance of catalysts in a cost-effective and time-efficient manner.

The third component emphasizes the importance of feedback loops, integrating real-time industrial data into AI models for continuous improvement. Adaptive learning systems can update predictive models based on new data, ensuring that the framework remains accurate and relevant as more information becomes available.

The paper also explores potential applications of the framework in industries such as refining and pharmaceutical manufacturing. By presenting case studies where AI-driven catalyst design could enhance performance by up to 50%, the paper highlights the practical benefits of adopting Catalysis 4.0.

2. Data-Driven Material Discovery

2.1. AI Algorithms and Material Properties

Artificial Intelligence has become a transformative force in numerous scientific domains, including the field of catalyst development. One of the most compelling applications of AI in this context is its ability to predict the performance of catalytic materials based on their intrinsic properties. AI algorithms, particularly machine learning models, are designed to identify complex patterns within large datasets, making them invaluable tools for understanding the intricate relationships between material properties and catalytic performance (Islam, Islam, Hossain Uzir, Abd Wahab, & Abdul Latiff, 2020).

Traditional catalyst development relies heavily on empirical testing, where researchers synthesize and test a vast array of materials to find those with optimal catalytic properties. This method is time-consuming, costly, and limited by human ability to discern subtle patterns within the data. AI algorithms, however, can process and analyze vast amounts of data far more efficiently (Lai et al., 2023). For instance, supervised learning algorithms can be trained on data about known catalysts and their performance metrics. Once trained, these models can predict the performance of new, untested materials, significantly narrowing down the number of candidates that need to be experimentally validated.

Moreover, AI can facilitate the discovery of novel catalysts with previously unexplored properties. Techniques such as neural networks and decision trees can uncover non-linear relationships between variables that traditional statistical methods might miss. This ability to delve deeper into the data can lead to breakthroughs in identifying materials with superior catalytic properties, accelerating innovation in this field (Salehmin et al., 2024).

2.2. Data Collection and Management

The effectiveness of AI algorithms in catalyst development depends on the quality and comprehensiveness of the data on which they are trained. High-quality data is critical for building reliable predictive models that can accurately forecast catalytic performance. Therefore, data collection and management are fundamental components of a successful AI-driven material discovery framework.

Data collection in this context involves gathering detailed information about various materials, including their chemical compositions, structural properties, and performance metrics in different catalytic reactions. This data can be sourced from experimental studies, published literature, and high-throughput screening experiments. High-throughput experimentation, in particular, generates vast amounts of data by testing thousands of material combinations in parallel, providing a rich dataset for AI analysis (Rodrigues, Florea, de Oliveira, Diamond, & Oliveira, 2021).

Managing this data effectively requires robust data infrastructure and standardized protocols. Data should be stored in a centralized database with standardized formats to ensure consistency and accessibility. Metadata, including information about the experimental conditions and methods used, should also be recorded to provide context for the data. Data preprocessing steps, such as normalization and error correction, are essential to ensure the data's accuracy and reliability (Habib & Okayli, 2024).

Effective data management also involves data integration, combining information from different sources to create a comprehensive dataset. This can involve data fusion and mining techniques to extract relevant information from disparate datasets. Researchers can build more robust models that account for a wider range of variables and conditions by integrating data from various sources (Krishnamurthi, Kumar, Gopinathan, Nayyar, & Qureshi, 2020).

2.3. Predictive Modeling

Predictive modeling is at the heart of AI-driven material discovery in catalyst development. Once high-quality data has been collected and managed, machine learning models can be employed to predict the performance of new catalytic materials. These models range from simple linear regressions to complex neural networks, each with its strengths and applications (Ninduwezuor-Ehiobu et al., 2023).

Linear regression models are often used as a starting point for predictive modeling. They are relatively simple to implement and can provide insights into the linear relationships between material properties and catalytic performance. However, many catalytic processes involve complex, non-linear interactions that linear models cannot capture (Kerner, Dogan, & von Recum, 2021).

For these more complex relationships, non-linear models such as decision trees, support vector machines (SVMs), and neural networks are more suitable. Decision trees are intuitive and can handle both numerical and categorical data, making them versatile tools for material discovery. Support vector machines are effective in high-dimensional spaces and can model non-linear relationships through the use of kernel functions (Mienye & Jere, 2024).

Neural networks, particularly deep learning models, have shown great promise in material science due to their ability to model highly complex, non-linear interactions. These models consist of multiple layers of interconnected neurons, which can learn hierarchical representations of the data. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are specialized types of neural networks that are particularly useful for image and sequence data, respectively. In catalyst development, CNNs can be used to analyze structural images of materials, while RNNs can model sequential data such as time-series performance metrics (Tang, Kurths, Lin, Ott, & Kocarev, 2020).

Training these models involves feeding them large amounts of data and adjusting their parameters to minimize prediction error. This process, known as supervised learning, requires a labeled dataset where the inputs (material properties) are paired with known outputs (catalytic performance). The model iteratively adjusts its parameters based on the error of its predictions, gradually improving its accuracy. Once trained, these predictive models can be used to screen new materials, forecast their performance and identify promising candidates for experimental validation. This approach drastically reduces the number of materials that need to be synthesized and tested, saving time and resources.

3. Virtual Testing Environment

3.1. Simulation Platform

The advent of advanced computational technologies has paved the way for the development of sophisticated simulation platforms that can revolutionize the field of catalyst development. A virtual testing environment for catalysts represents a significant leap forward, enabling researchers to evaluate and optimize catalytic materials without the need for extensive physical experimentation. This proposed simulation platform leverages the power of computational modeling and machine learning algorithms to predict the behavior and performance of catalysts under various conditions (Isahak & Al-Amiery, 2024).

At the core of this simulation platform is a robust computational framework that integrates multiple modeling techniques, including molecular dynamics (MD), density functional theory (DFT), and kinetic Monte Carlo (kMC) simulations. These techniques allow for the detailed examination of catalytic processes at different scales, from the atomic level to macroscopic reactions. Researchers can gain insights into the fundamental mechanisms that drive catalytic activity by simulating the interactions between atoms and molecules (Cheimarios, To, Kokkoris, Memos, & Boudouvis, 2021).

The platform is designed to be user-friendly, with an intuitive interface that allows researchers to input the properties of catalytic materials and specify the conditions under which they wish to test them. The system then uses advanced algorithms to simulate the catalytic reactions, providing detailed outputs on reaction rates, selectivity, and stability. These outputs can be used to compare different materials and identify those with the most promising characteristics for further development.

3.2. Industrial Conditions Simulation

One of the critical aspects of a virtual testing environment is its ability to simulate industrial conditions accurately. Catalysts often operate under extreme conditions, such as high temperatures, pressures, and varying chemical environments, which can significantly impact their performance. Therefore, it is essential that the simulation platform can replicate these conditions to provide realistic assessments of catalyst efficiency.

The simulation platform incorporates models that mimic industrial processes' thermodynamic and kinetic parameters to achieve this. For instance, molecular dynamics simulations can be used to study the behavior of catalysts at high temperatures, providing insights into thermal stability and reaction kinetics. Similarly, pressure effects can be modeled using equations of state that describe how catalysts interact with gases and liquids under different pressures (Liu et al., 2023).

Additionally, the platform includes modules for simulating the presence of impurities and byproducts that are common in industrial processes. These simulations can help identify potential deactivation pathways and inform the design of more robust catalysts. By incorporating these industrial conditions into the simulation framework, researchers can ensure that the virtual testing environment comprehensively evaluates catalyst performance (Atsonios et al., 2020).

3.3. Benefits of Virtual Testing

The use of virtual testing environments offers several significant advantages over traditional experimental methods. First and foremost, simulations can drastically reduce the time and cost associated with catalyst development. Traditional methods require synthesizing and testing numerous candidate materials, each involving significant resource investment. In contrast, virtual testing allows researchers to screen a vast number of materials computationally, identifying the most promising candidates before any physical experimentation takes place.

Moreover, simulations provide a level of detail and control that is often unattainable in physical experiments. Researchers can precisely manipulate individual variables, such as temperature or pressure, and observe their direct

impact on catalytic performance. This ability to conduct controlled studies helps in understanding the fundamental principles of catalysis and aids in the rational design of new materials (Chen & Wang, 2022).

Another key benefit is the ability to explore extreme conditions and scenarios that might be impractical or unsafe to replicate in a laboratory setting. For example, testing catalysts at very high temperatures or pressures can pose significant safety risks and logistical challenges. Virtual testing eliminates these concerns, allowing researchers to investigate the limits of catalyst performance safely (Cornish et al., 2021).

Furthermore, virtual testing environments can facilitate the integration of machine learning models, enhancing the predictive power of simulations. Machine learning algorithms can be trained on simulation data to identify patterns and predict the behavior of new materials. This iterative process of simulation and machine learning creates a feedback loop that continuously improves the accuracy and reliability of the virtual testing environment (Chan, Van Gerven, Dubois, & Bernaerts, 2021).

The insights gained from virtual testing can also accelerate the commercialization of new catalysts. By providing detailed performance data early in development, simulations can help identify the most viable candidates for scale-up and industrial application. This can shorten the time to market for new catalytic materials, driving innovation and competitiveness in various industries.

4. Feedback Loops and Continuous Improvement

4.1. Real-Time Data Integration

The integration of real-time industrial data into AI models represents a significant advancement in catalyst development. Real-time data integration ensures that AI models are continually updated with the latest information, allowing for more accurate predictions and timely adjustments. In the context of catalyst development, this means that performance metrics, environmental conditions, and reaction outcomes can be continuously fed into the AI system, providing a dynamic and up-to-date dataset (Fisher et al., 2020).

One of the primary benefits of real-time data integration is its ability to capture the variability and complexity of industrial processes. Catalysts operate under diverse and often fluctuating conditions, which can impact their performance in unpredictable ways. By incorporating real-time data, AI models can account for these variations, making their predictions more robust and reliable. For instance, if a catalyst's performance deteriorates due to an unforeseen impurity in the feedstock, real-time data can help identify this issue promptly, allowing for rapid adjustments (Udugama et al., 2020).

Furthermore, real-time data integration enhances the ability of AI models to perform predictive maintenance and optimize operational parameters. By continuously monitoring catalyst performance, AI systems can detect early signs of degradation or inefficiency, prompting preemptive measures to mitigate potential issues. This proactive approach extends the lifespan of catalysts and ensures consistent and optimal performance, leading to significant cost savings and operational efficiencies (Arinze, Izionworu, Isong, Daudu, & Adefemi, 2024).

4.2. Adaptive Learning Systems

Adaptive learning systems are crucial to continuous improvement in AI-driven catalyst development. These systems are designed to evolve and improve over time by learning from new data and experiences. In essence, adaptive learning systems enable AI models to refine their predictions and recommendations as they encounter new information, creating a cycle of perpetual enhancement.

One of the key mechanisms behind adaptive learning is the use of feedback loops. Feedback loops involve the continuous feeding of new data back into the AI model, allowing it to update its parameters and improve its accuracy. In catalyst development, this means that every new experiment, observation, or industrial application provides valuable data that can enhance the AI's predictive capabilities. For example, suppose a new catalyst formulation shows unexpected performance characteristics. In that case, the data from this experiment can be fed back into the AI model, helping it to learn and adapt its predictions for future formulations (Sjödin, Parida, Palmié, & Wincent, 2021).

Another critical aspect of adaptive learning systems is their ability to perform real-time analysis and adjustment. This involves updating model parameters and recalibrating the models to account for changes in the operating environment. For instance, if a catalyst is used in a process with fluctuating temperatures and pressures, the adaptive learning system

can continuously adjust its predictions and recommendations to reflect these changes, ensuring optimal performance under varying conditions.

The integration of adaptive learning systems also facilitates the development of more sophisticated AI models that can handle complex, multi-dimensional data. Advanced machine learning techniques, such as reinforcement learning and neural networks, are particularly well-suited for adaptive learning. These techniques enable AI models to learn from their mistakes and successes, gradually improving their performance. In the context of catalyst development, this means that AI models can become increasingly adept at predicting catalytic performance, identifying promising materials, and optimizing operational parameters (Bauer, Heigl, Hinz, & Kosfeld, 2024).

4.3. Impact on Catalyst Development Cycle

The incorporation of feedback loops and adaptive learning systems has a profound impact on the speed and efficiency of the catalyst development cycle. Traditionally, the development of new catalysts has been a lengthy and iterative process, involving extensive trial-and-error experimentation. However, this process can be significantly accelerated with the advent of AI-driven methodologies. One of the most significant impacts of feedback loops is the reduction in development time. AI models can rapidly identify promising catalysts and optimize their performance by continuously integrating real-time data and learning from new experiments. This iterative data integration and model refinement process allows researchers to make informed decisions more quickly, reducing the time required to develop new catalysts from years to months or weeks (Lai et al., 2023).

Moreover, feedback loops enhance the efficiency of the development process by minimizing the need for redundant experiments. Researchers often have to conduct numerous experiments in traditional catalyst development to explore different material combinations and conditions. With AI-driven feedback loops, much of this experimentation can be simulated virtually, with the AI model predicting the most promising candidates for physical testing. This targeted approach reduces the number of experiments required, saving both time and resources.

The continuous improvement facilitated by adaptive learning systems also leads to developing higher-quality catalysts. As AI models learn from each iteration, they better predict performance and identify potential issues. This means that the catalysts developed through this process are likely to be more effective and reliable, with fewer performance inconsistencies. The ability to continuously refine and optimize catalysts ensures that the materials produced are of the highest possible quality, meeting the rigorous demands of industrial applications.

Furthermore, the integration of real-time data and adaptive learning systems fosters a culture of innovation and agility within the catalyst development field. Researchers are empowered to experiment with new ideas and approaches, confident that their AI tools can quickly assess and refine these innovations. This dynamic environment encourages the exploration of novel materials and techniques, driving continuous improvement and technological advancement.

5. Conclusion

This paper has outlined a comprehensive framework for integrating artificial intelligence and machine learning into the field of catalyst development, termed Catalysis 4.0. The framework is built on three primary components: data-driven material discovery, virtual testing environments, and feedback loops for continuous improvement. The data-driven material discovery phase leverages AI algorithms to predict catalytic performance based on extensive material properties and high-quality data. The virtual testing environment provides a simulation platform to evaluate catalyst efficiency under various industrial conditions, significantly reducing the time and cost associated with traditional experimentation. Finally, feedback loops and adaptive learning systems ensure continuous refinement and enhancement of AI models by integrating real-time industrial data, thereby accelerating the catalyst development cycle and improving overall efficiency.

The integration of AI and ML into catalyst development holds transformative potential across various industries. In refining, for instance, quickly identifying and optimizing new catalytic materials can lead to more efficient conversion processes, reducing energy consumption and operational costs. Enhanced catalysts can also improve the selectivity and yield of valuable products, directly impacting profitability. In pharmaceutical manufacturing, where precision and reliability are paramount, AI-driven catalysts can streamline the synthesis of complex molecules, enhancing reaction efficiency and reducing the occurrence of undesirable by-products. This expedites the drug development process and ensures higher purity and consistency in pharmaceutical products.

Moreover, the framework's impact extends beyond refining and pharmaceuticals. Industries such as petrochemicals, environmental engineering, and renewable energy can benefit from advanced catalysts' rapid and cost-effective development. For example, in environmental engineering, improved catalysts can enhance the efficiency of pollution control technologies, contributing to cleaner air and water. In renewable energy, catalysts play a critical role in hydrogen production and carbon capture processes. AI-driven advancements in these areas can significantly boost the adoption and effectiveness of sustainable technologies.

Recommendations

To fully realize the potential of AI and ML in catalyst development, further research and development are essential. One key area for future exploration is the enhancement of data collection and management techniques. High-quality, comprehensive datasets are crucial for training robust AI models. Collaborative efforts across academia, industry, and government agencies can facilitate the creation of standardized databases that encompass a wide range of material properties and performance metrics.

Another important direction is the development of more sophisticated AI algorithms that can handle the complexity and variability of catalytic processes. Techniques such as reinforcement learning, which allows models to learn from their own successes and failures, can be particularly valuable. Additionally, the integration of multi-omics data—encompassing genomics, proteomics, and metabolomics—into AI models can provide deeper insights into the molecular mechanisms underlying catalysis, leading to the discovery of novel catalysts with unprecedented performance characteristics.

Furthermore, fostering interdisciplinary collaboration between material scientists, chemists, and AI experts is critical. Such collaborations can bridge the gap between theoretical research and practical application, ensuring that AI-driven catalysts are innovative and viable for industrial deployment. Establishing platforms for knowledge exchange and joint research initiatives can accelerate the development and implementation of Catalysis 4.0.

Finally, addressing AI-driven catalyst development's ethical and environmental implications is paramount. Ensuring that the advancements in this field contribute to sustainability and do not inadvertently harm the environment is essential. Developing guidelines and best practices for the responsible use of AI in material science can help mitigate potential risks and promote the long-term benefits of these technologies.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Al-Akayleh, F., Ali Agha, A. S., Abdel Rahem, R. A., & Al-Remawi, M. (2024). A mini review on the applications of artificial intelligence (AI) in surface chemistry and catalysis. *Tenside Surfactants Detergents*(0).

[2] Arinze, C. A., Izionworu, V. O., Isong, D., Daudu, C. D., & Adefemi, A. (2024). Integrating artificial intelligence into engineering processes for improved efficiency and safety in oil and gas operations. *Open Access Research Journal of Engineering and Technology, 6*(1), 39-51.

[3] Atsonios, K., Nesiadis, A., Detsios, N., Koutita, K., Nikolopoulos, N., & Grammelis, P. (2020). Review on dynamic process modeling of gasification based biorefineries and bio-based heat & power plants. *Fuel Processing Technology, 197*, 106188.

[4] Bauer, K., Heigl, R., Hinz, O., & Kosfeld, M. (2024). Feedback loops in machine learning: A study on the interplay of continuous updating and human discrimination. *Journal of the Association for Information Systems, 25*(4), 804-866.

[5] Benavides-Hernández, J., & Dumeignil, F. (2024). From Characterization to Discovery: Artificial Intelligence, Machine Learning and High-Throughput Experiments for Heterogeneous Catalyst Design. *ACS Catalysis, 14*(15), 11749-11779.

[6] Chan, P., Van Gerven, T., Dubois, J.-L., & Bernaerts, K. (2021). Virtual chemical laboratories: A systematic literature review of research, technologies and instructional design. *Computers and Education Open, 2*, 100053.

[7] Cheimarios, N., To, D., Kokkoris, G., Memos, G., & Boudouvis, A. G. (2021). Monte Carlo and kinetic Monte Carlo models for deposition processes: a review of recent works. *Frontiers in Physics, 9*, 631918.

[8] Chen, J., & Wang, L. (2022). Effects of the catalyst dynamic changes and influence of the reaction environment on the performance of electrochemical CO2 reduction. *Advanced Materials, 34*(25), 2103900.

[9] Cornish, N. E., Anderson, N. L., Arambula, D. G., Arduino, M. J., Bryan, A., Burton, N. C., . . . Griffith, N. K. (2021). Clinical laboratory biosafety gaps: lessons learned from past outbreaks reveal a path to a safer future. *Clinical microbiology reviews, 34*(3), 10.1128/cmr. 00126-00118.

[10] Fisher, O. J., Watson, N. J., Escrig, J. E., Witt, R., Porcu, L., Bacon, D., . . . Gomes, R. L. (2020). Considerations, challenges and opportunities when developing data-driven models for process manufacturing systems. *Computers & Chemical Engineering, 140*, 106881.

[11] Habib, M., & Okayli, M. (2024). Evaluating the sensitivity of machine learning models to data preprocessing technique in concrete compressive strength estimation. *Arabian Journal for Science and Engineering*, 1-19.

[12] Hughes, A. E., Haque, N., Northey, S. A., & Giddey, S. (2021). Platinum group metals: A review of resources, production and usage with a focus on catalysts. *Resources, 10*(9), 93.

[13] Isahak, W. N. R. W., & Al-Amiery, A. (2024). Catalysts driving efficiency and innovation in thermal reactions: A comprehensive review. *Green Technologies and Sustainability*, 100078.

[14] Islam, A., Islam, M., Hossain Uzir, M., Abd Wahab, S., & Abdul Latiff, A. (2020). The panorama between COVID-19 pandemic and Artificial Intelligence (AI): Can it be the catalyst for Society 5.0. *International Journal of Scientific Research and Management, 8*(12), 2011-2025.

[15] Johnson, B. (2022). *Making ammonia: fritz haber, walther nernst, and the nature of scientific discovery*: Springer Nature.

[16] Kerner, J., Dogan, A., & von Recum, H. (2021). Machine learning and big data provide crucial insight for future biomaterials discovery and research. *Acta Biomaterialia, 130*, 54-65.

[17] Krishnamurthi, R., Kumar, A., Gopinathan, D., Nayyar, A., & Qureshi, B. (2020). An overview of IoT sensor data processing, fusion, and analysis techniques. *Sensors, 20*(21), 6076.

[18] Lai, N. S., Tew, Y. S., Zhong, X., Yin, J., Li, J., Yan, B., & Wang, X. (2023). Artificial intelligence (AI) workflow for catalyst design and optimization. *Industrial & Engineering Chemistry Research, 62*(43), 17835-17848.

[19] Liu, S., Wei, L., Zhou, Q., Yang, T., Li, S., & Zhou, Q. (2023). Simulation strategies for ReaxFF molecular dynamics in coal pyrolysis applications: A review. *Journal of Analytical and Applied Pyrolysis, 170*, 105882.

[20] Mienye, I. D., & Jere, N. (2024). A Survey of Decision Trees: Concepts, Algorithms, and Applications. *IEEE Access*.

[21] Ninduwezuor-Ehiobu, N., Tula, O. A., Daraojimba, C., Ofonagoro, K. A., Ogunjobi, O. A., Gidiagba, J. O., . . . Banso, A. A. (2023). Tracing the evolution of ai and machine learning applications in advancing materials discovery and production processes. *Engineering Science & Technology Journal, 4*(3), 66-83.

[22] Papadimitriou, I., Gialampoukidis, I., Vrochidis, S., & Kompatsiaris, I. (2024). AI methods in materials design, discovery and manufacturing: A review. *Computational Materials Science, 235*, 112793.

[23] Rane, N., Kaya, O., & Rane, J. (2024). Artificial intelligence, machine learning, and deep learning technologies as catalysts for industry 4.0, 5.0, and society 5.0. *Artificial Intelligence, Machine Learning, and Deep Learning for Sustainable Industry, 5*, 2-2.

[24] Rodrigues, J. F., Florea, L., de Oliveira, M. C., Diamond, D., & Oliveira, O. N. (2021). Big data and machine learning for materials science. *Discover Materials, 1*, 1-27.

[25] Rubtsov, N. M., Troshin, K. Y., & Alymov, M. I. (2023). *Catalytic Ignition of Hydrogen and Hydrogen-Hydrocarbon Blends Over Noble Metals*: Springer.

[26] Sahai, A. K., & Rath, N. (2021). Artificial intelligence and the 4th industrial revolution. In *Artificial intelligence and machine learning in business management* (pp. 127-143): CRC Press.

[27] Salehmin, M. N. I., Kiong, T. S., Mohamed, H., Umar, D. A., Yu, K. L., Ong, H. C., . . . Lim, S. S. (2024). Navigating challenges and opportunities of machine learning in hydrogen catalysis and production processes: Beyond algorithm development. *Journal of Energy Chemistry*.

[28] Sjödin, D., Parida, V., Palmié, M., & Wincent, J. (2021). How AI capabilities enable business model innovation: Scaling AI through co-evolutionary processes and feedback loops. *Journal of Business Research, 134*, 574-587.

[29] Suwardi, A., Wang, F., Xue, K., Han, M. Y., Teo, P., Wang, P., . . . Li, Z. (2022). Machine learning-driven biomaterials evolution. *Advanced Materials, 34*(1), 2102703.

[30] Tang, Y., Kurths, J., Lin, W., Ott, E., & Kocarev, L. (2020). Introduction to focus issue: When machine learning meets complex systems: Networks, chaos, and nonlinear dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science, 30*(6).

[31] Udugama, I. A., Gargalo, C. L., Yamashita, Y., Taube, M. A., Palazoglu, A., Young, B. R., . . . Bayer, C. (2020). The role of big data in industrial (bio) chemical process operations. *Industrial & Engineering Chemistry Research, 59*(34), 15283-15297.