I&EC research
Industrial & Engineering Chemistry Research
pubs.acs.org/IECR
Editorial

Machine Learning and Data Science in Chemical Engineering

Cite This: *Ind. Eng. Chem. Res.* 2022, 61, 8357−8358
Read Online
ACCESS | Metrics & More | Article Recommendations

Downloaded via UNIV OF PITTSBURGH on June 24, 2023 at 22:13:49 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Chemical engineering is a data-rich subject. Practitioners collect and analyze data for understanding flow patterns, developing empirical models, designing and optimizing chemical reactions, as well as monitoring and controlling chemical processes and systems. Because of the high data dependence, it seems natural that artificial intelligence and machine learning would play an important role in the field of chemical engineering. However, for a long time in the 20th century, because of the lack in computing power and infrastructure of data generation and storage, limited success was observed.[1,2] Recent years have seen a surge in machine learning related research from the computer science discipline to many exciting interdisciplinary areas, including chemistry, chemical engineering, and biological engineering.[3−11] This trend creates opportunities for a transformative paradigm in our chemical engineering community. Furthermore, this exciting development could be attributed to both rapid advancements in computer hardware and algorithms that enable the training of large and complex models, advanced simulation and experimentation tools, and systems that accelerate the accumulation of data as well as better practices for data documentation and management. Researchers are now equipped with powerful tools in data science and machine learning to tackle complex challenges and opportunities, both at macroscopic and microscopic levels, facing the disciplines of chemical and biomolecular engineering and the relevant industries.

One primary usage of machine learning methods is to approximate high complex and nonlinear relationships between input and output variables. This type of problem is ubiquitous in chemical engineering, including but not limited to the prediction of molecular properties from structures (especially for multicomponent systems), development of closure models for flow dynamics and transport, reconstruction and identification of flow patterns, quantification of uncertainty in physical models, prediction and optimization of reaction outcome and reactor performance based on the reaction conditions (e.g., reagents, catalysts, solvents, reactants, products, flow rate, temperature, and pressure), assistance in building better catalyst models, understanding catalytic mechanisms, prediction of process response to control actions, etc. The application of machine learning, especially deep learning methods, provides an effective solution to these problems, because of the large parameter space to account for complex functions. In addition, flexible model architectures can be designed to process complicated information. For example, convolutional neural networks capture local connections and can be used to process data with spatial structures, such as images and molecular graphs. Recurrent neural networks account for sequential connections and are used for time series data. The physics-informed machine learning that embeds the prior physics, reaction mechanisms, or other types of constraints into its architectures can greatly boost interpretability and approximation ability of pure machine learning. These methods are increasingly helpful as more complex data are generated and used in chemical engineering developments, including images of flow patterns from high-speed cameras, flow field data from high-fidelity numerical simulations, images of molecular structures from electronic microscopes, time series data from process controllers, etc. Finally, machine learning frameworks such as active learning and reinforcement learning, can be used to aid decision making processes in chemical engineering, including iterative experimental design, process scheduling, control, and optimization.

This special issue of *Industrial & Engineering Chemistry Research* presents an excellent collection of articles from internationally renowned researchers from all around the world to showcase the application of machine learning and data science in the aforementioned chemical engineering problems.[12−30] We truly appreciate the efforts from all contributing authors to make it happen. We hope these articles provide new insights and perspectives as to how machine learning can be used in a wide variety of chemical engineering problems, and stimulate more creative solutions to existing and future challenges.

Hanyu Gao orcid.org/0000-0002-6346-0739
Li-Tao Zhu orcid.org/0000-0001-6514-8864
Zheng-Hong Luo orcid.org/0000-0001-9011-6020
Marco A. Fraga orcid.org/0000-0002-9768-8360
I-Ming Hsing

AUTHOR INFORMATION

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.iecr.2c01788

Notes

Views expressed in this editorial are those of the authors and not necessarily the views of the ACS.

Special Issue: Machine Learning and Data Science in Chemical Engineering

Published: June 22, 2022

REFERENCES

(1) Hattori, T.; Kito, S. Neural Network as a Tool for Catalyst Development. *Catal. Today* **1995**, *23* (4), 347−355.

(2) Dobbelaere, M. R.; Plehiers, P. P.; Van de Vijver, R.; Stevens, C. V; Van Geem, K. M. Machine Learning in Chemical Engineering: Strengths, Weaknesses, Opportunities, and Threats. *Engineering* **2021**, *7*, 1201−1211.

(3) Coley, C. W.; Green, W. H.; Jensen, K. F. Machine Learning in Computer-Aided Synthesis Planning. *Acc. Chem. Res.* **2018**, *51* (5), 1281−1289.

(4) Coley, C. W.; Thomas, D. A.; Lummiss, J. A. M.; Jaworski, J. N.; Breen, C. P.; Schultz, V.; Hart, T.; Fishman, J. S.; Rogers, L.; Gao, H.; et al. A Robotic Platform for Flow Synthesis of Organic Compounds Informed by AI Planning. *Science (80)* **2019**, *365* (6453), eaax1566.

(5) Gao, H.; Struble, T. J.; Coley, C. W.; Wang, Y.; Green, W. H.; Jensen, K. F. Using Machine Learning to Predict Suitable Conditions for Organic Reactions. *ACS Central Sci.* **2018**, *4* (11), 1465−1476.

(6) Segler, M. H. S.; Preuss, M.; Waller, M. P. Planning Chemical Syntheses with Deep Neural Networks and Symbolic AI. *Nature* **2018**, *555* (7698), 604.

(7) Schwaller, P.; Petraglia, R.; Zullo, V.; Nair, V. H.; Haeuselmann, R. A.; Pisoni, R.; Bekas, C.; Iuliano, A.; Laino, T. Predicting Retrosynthetic Pathways Using Transformer-Based Models and a Hyper-Graph Exploration Strategy. *Chem. Sci.* **2020**, *11* (12), 3316−3325.

(8) Guo, J.; Ibanez-Lopez, A. S.; Gao, H.; Quach, V.; Coley, C. W.; Jensen, K. F.; Barzilay, R. Automated Chemical Reaction Extraction from Scientific Literature. *J. Chem. Inf. Model.* **2022**, *62*, 2035−2045.

(9) Vaucher, A. C.; Zipoli, F.; Geluykens, J.; Nair, V. H.; Schwaller, P.; Laino, T. Automated Extraction of Chemical Synthesis Actions from Experimental Procedures. *Nat. Commun.* **2020**, *11* (1), 3601.

(10) Schwaller, P.; Gaudin, T.; Lanyi, D.; Bekas, C.; Laino, T. Found in Translation": Predicting Outcomes of Complex Organic Chemistry Reactions Using Neural Sequence-to-Sequence Models. *Chem. Sci.* **2018**, *9* (28), 6091−6098.

(11) Wang, X.; Qian, Y.; Gao, H.; Coley, C. W.; Mo, Y.; Barzilay, R.; Jensen, K. F. Towards Efficient Discovery of Green Synthesis Pathways with Monte Carlo Tree Search and Reinforcement Learning. *Chem. Sci.* **2020**, *11*, 10959−10972.

(12) Nishimura, S.; Ohyama, J.; Li, X.; Miyazato, I.; Taniike, T.; Takahashi, K. Machine Learning-Aided Catalyst Modification in Oxidative Coupling of Methane via Manganese Promoter. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c05079.

(13) Zhang, M.; Cao, D.; Lan, X.; Shi, X.; Gao, J. An Ensemble-Learning Approach To Predict the Coke Yield of Commercial FCC Unit. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04735.

(14) Doval, F.; Soares, R.; Secchi, A. R.; Bezerra De Souza, M. Development of a Nonlinear Model Predictive Control for Stabilization of a Gas-Lift Oil Well. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04728.

(15) Tanudjaja, H. J.; Chew, J. W. Application of Machine Learning-Based Models to Understand and Predict Critical Flux of Oil-in-Water Emulsion in Crossflow Microfiltration. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c046.

(16) Alizadeh, R.; Mohebbi Najm Abad, J.; Fattahi, A.; Mesgarpour, M.; Doranehgard, M. H.; Xiong, Q.; Karimi, N. Machine-Learning Enhanced Analysis of Mixed Biothermal Convection of Single Particle and Hybrid Nanofluids within a Complex Configuration. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04662.

(17) Li, Y.; Wang, Y.; Chen, Y.; Lu, Y.; Hua, K.; Ren, J.; Mozafari, G.; Lu, Q.; Cao, Y. Deep-Learning-Based Predictive Control of Battery Management for Frequency Regulation. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04903.

(18) Ong, L.; Karmakar, G.; Atherton, J.; Zhou, X.; Lim, M. Q.; Chadzynski, A.; Li, L.; Wang, X.; Kraft, M. Embedding Energy Storage Systems into a Dynamic Knowledge Graph. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c03838.

(19) Marcato, A.; Boccardo, G.; Marchisio, D. From Computational Fluid Dynamics to Structure Interpretation via Neural Networks: An Application to Flow and Transport in Porous Media. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04760.

(20) Wang, J.; Dong, T.; Cheng, Y.; Yan, W.-C. Machine Learning Assisted Spraying Pattern Recognition for Electrohydrodynamic Atomization System. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04669.

(21) Dobbelaere, M. R.; Ureel, Y.; Vermeire, F. H.; Tomme, L.; Stevens, C. V.; Van Geem, K. M. Machine Learning for Physicochemical Property Prediction of Complex Hydrocarbon Mixtures. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.2c00442.

(22) Campos, G.; El-Farra, N. H.; Palazoglu, A. Soft Actor-Critic Deep Reinforcement Learning with Hybrid Mixed-Integer Actions for Demand Responsive Scheduling of Energy Systems. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04984.

(23) Zhan, N.; Kitchin, J. R. Model-Specific to Model-General Uncertainty for Physical Properties. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04706.

(24) Ma, L.; Kashanj, S.; Xu, S.; Zhou, J.; Nobes, D. S.; Ye, M. Flow Reconstruction and Prediction Based on Small Particle Image Velocimetry Experimental Datasets with Convolutional Neural Networks. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04704.

(25) Luo, J.; Canuso, V.; Jang, J. B.; Wu, Z.; Morales-Guio, C. G.; Christofides, P. D. Machine Learning-Based Operational Modeling of an Electrochemical Reactor: Handling Data Variability and Improving Empirical Models. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04176

(26) Yang, S. D.; Ali, Z. A.; Kwon, H.; Wong, B. M. Predicting Complex Erosion Profiles in Steam Distribution Headers with Convolutional and Recurrent Neural Networks. *Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.1c04712.

(27) Zhao, X.; Luo, T.; Jin, H. Predicting Diffusion Coefficients of Binary and Ternary Supercritical Water Mixtures via Machine and Transfer Learning with Deep Neural Network. *Ind. Eng. Chem. Res.* **2022**. DOI: 10.1021/acs.iecr.2c00017.

(28) Xie, Z.; Gu, X.; Shen, Y. A Machine Learning Study of Predicting Mixing and Segregation Behaviors in a Bidisperse Solid-Liquid Fluidized Bed. *Ind. Eng. Chem. Res.* **2022**. DOI: 10.1021/acs.iecr.2c00071.

(29) Shi, Y.; Wang, J.; Wang, Q.; Jia, Q.; Yan, F.; Luo, Z. H.; Zhou, Y. N. Supervised Machine Learning Algorithms for Predicting Rate Constants of Ozone Reaction with Micropollutants. *Ind. Eng. Chem. Res.* **2022**. DOI: 10.1021/acs.iecr.1c04697.

(30) Zipoli, F.; Viterbo, V.; Schliter, O.; Kahle, L.; Laino, T. Prediction of Phase Diagrams and Associated Phase Structural Properties.*Ind. Eng. Chem. Res.* **2022**, DOI: 10.1021/acs.iecr.2c00355.

https://doi.org/10.1021/acs.iecr.2c01788
Ind. Eng. Chem. Res. 2022, 61, 8357−8358