***Artificial Intelligence applied to Chemical Engineering***

Luigi Piero Di Bonito1,2\*, Lelio Campanile1, Mauro Iacono1, Francesco Di Natale2

*1 Dipartimento di Matematica e Fisica – Università degli Studi della Campania “Luigi Vanvitelli”*

*2 Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione Industriale – Università degli Studi di Napoli “Federico II”*

*\*Corresponding author E-Mail: luigipiero.dibonito@unicampania.it*

**1.Introduction**

Chemical engineers have found mathematical modeling to be extremely useful in understanding and creating chemical processes. However, for real systems, many of these models cannot be solved analytically and require a significant amount of processing power to for a numerical resolution. The transition towards a green and sustainable chemical industry necessitates the development of new design paradigms for flexible plants, smart molecules, and functional materials. Decades of modeling, simulations, experimental tests, and sensors deployed on chemical plants have generated huge amounts of data for Chemical Engineering (CE) community [1], adding the option of generating predictions based on experience, as an additional modeling toolbox for specialist designers. Machine Learning (ML) models are statistical and mathematical models that can "learn" from experience and find patterns in data without requiring explicit, rule-based programming. Machine learning including Deep Learning (DL) and Artificial Intelligence (AI), has the potential to overcome the constraints of mechanistic modeling: ML approaches can learn complicated behavior, model building is inexpensive, and optimization is possible. Moreover they allow modeling the structures governing the processes that generate the large data-collections, with problem-independent approaches.

Diagram, schematic

Description automatically generated

**Figure 1.** An example of raw materials: big data from chemical plants

This work is part of a new Italian Ph.D. Programme (funded by the PON Ricerca e Innovazione 2014-2020) aimed to development of new AI assisted methodologies for the design of chemical plants. In this paper we will present an overview of AI/ML approaches applied to CE, demonstrating how these techniques offer unique perspectives, but that only the integration between AI and CE groups can fully realize their promise.

**2. Methods**

The purpose of this work is an overview of the actual state of the art on ML/DL/AI methodologies used in the field of plant design and operation. The number of papers related to the application of ML/DL/AI on chemical plants has grown exponentially in the last ten years. The goal is to identify the primary study fields and applications in which they have been employed, as well as the approaches and datasets that have been collected, reorganized, and used. To achieve our goals, we explored the literature finding some paradigmatic references and our research work has been driven by the following research questions:

I. According to the literature, what types of issues have been solved?

II. What ML/DL/AI approaches are employed in these works?

III. What sort of dataset has been used?

**3. Results and discussion**

We decided to evaluate several aspects of this research subject, primarily to demonstrate to the chemical engineering community the potential of ML/DL/AI approaches applied to chemical processes. To succeed this goal, and in response to the first question (I), we investigated the types of issues solved using these strategies. To the best of our knowledge, there are no example of ML/DL/AI methods used for plant design, and the literature survey revealed that the majority of ML/DL/AI applications to chemical plants are related to prediction, optimization, fault diagnosis, and control issues.

For example, in terms of prediction/modeling challenges, AI has been used to predict NOx emissions from coal-powder power plants [2] and the operation of a wet scrubber system for air pollution management [3]. In terms of optimization problems, Osuolale and Zhang [4] proposed a methodology for optimizing the energy efficiency of an atmospheric distillation unit without sacrificing product quality or process throughput, whereas Zhang et al. [5] present an AI-based real time optimization (RTO) for two chemical process examples: a Continuous Stirred Tank Reactor (CSTR) and a distillation column. Concerning control issues, after the RTO analysis, the authors successfully updated the control systems of both processes using AI approaches [5]. Al-Dunainawi and Abbod [6] established an AI based control-logic system to regulate product compositions of distillation column.

As regards fault analysis, Mamandipoor et al. [7] monitored and analysed flows and compositions of the intermediate streams of a wastewater treatment plant, while Li et al. [8] proposed a fault diagnostic system for a distillation process.

Regarding the second question (II), a classification of ML/DL approaches, which comprise AI techniques is needed, into two broad categories: supervised learning and unsupervised learning. Supervised learning approaches use labeled data to train a model with an explicit input-output structure and learn functions that translate an input to an output. Unsupervised learning refers to a set of approaches that explore "unlabeled data", or data that lack a clear input-output relationship. Table 1 resumes the artificial intelligence approaches used in the previously listed references based on this classification.

**Table 1.** Overview of ML/DL/AI Applied Techniques

|  |  |  |  |
| --- | --- | --- | --- |
| **ML/DL/AI TECHNIQUES** | **SUPERVISED** | **UNSUPERVISED** | **REFERENCES** |
| Convolutional Neural Networks (CNN) | X |  | [2], [8] |
| Artificial Neural Networks (ANN) | X |  | [3], [5] |
| Bootstrap Aggregated Neural Networks (BNN) | X |  | [4] |
| Particle Swarm Optimization (PSO) | n.a. | n.a. | [6] |
| Recurrent Neural Networks (RNN) | X |  | [7] |
| Long-Short Term Memory Networks LSTM) | X |  | [7] |
| Deep Auto Encoders (DAE) |  | X | [8] |

To answer the third and final question (III), it is important to quickly describe the most common datasets within CE. Chemical engineering data sources may be divided into five primary categories: 1) data from the factory, 2) data from the laboratory, 3) data from simulation, 4) data from the literature, and 5) data from the company. According to this classification, Table 2 shows what type of dataset was used in prior references.

**Table 2**. Kind of Dataset used in AI Applications to CE

|  |  |
| --- | --- |
| **KIND OF DATASET** | **REFERENCES** |
| Plant Data | [2], [7] |
| Laboratory Data | [3] |
| Simulation Data | [4], [5], [8] |

**4. Conclusions**

In recent years, the literature in chemical engineering has mostly explored the integration of traditional analytical models and artificial neural networks for prediction/modeling, optimization, fault detection, and control concerns based on plant, laboratory, or simulation datasets. No examples of AI based plant design have been found.

The authors envisage that in the design of chemical plants, a hybrid approach has a great advantage respect to pure ML/DL/AI methods [9]. Indeed, the conventional analytical design models developed by chemical engineering since the last century constitute a corpus of knowledge in the form of chemical-physical, balances, transport, reaction kinetics and “engineering” equations, graphs and dataset which allow to design a given equipment (e.g., distillation towers, absorbers, mixers…) almost regardless of the specific chemical system at hand. This is the essence of the powerful concept of Unit Operations. Differently, a pure ML/DL/AI model may be exceptionally good in describing an existing process, but the design cannot be extended to different chemical systems. In a few words, its direct application to chemical plant design may push back chemical engineering development to the level of ancient alchemic recipes. Further efforts are thus needed to integrate the ML/DL/AI and conventional models to reach a better comprehension of plant design and operation, preserving the general schemes and design frameworks of Unit operations. A first example of a similar approach has been recently proposed by Napolitano et al. [10] for the case of a marine scrubber and in the review of McBride et al [9].

**References**

[1] R. H. Hariri, E. M. Fredericks, and K. M. Bowers, “Uncertainty in big data analytics: survey, opportunities, and challenges,” *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0206-3.

[2] I. Bukovský and M. Kolovratník, “A Neural Network Model for Predicting NOx at the Mělník 1,” *Acta Polytech.*, vol. 52, no. 3, 2012, doi: 10.14311/1538.

[3] B. A. Danzomo, M. Salami, and R. M. Khan, “Artificial Neural Network Model for Wet Scrubber Performance,” *Int. J. Sci. Eng. Res.*, vol. 3, no. 11, pp. 8–10, 2012, [Online]. Available: http://www.ijser.org/researchpaper%5CArtificial-Neural-Network-Model-for-Predicting-Wet-Scrubber-Performance.pdf.

[4] F. N. Osuolale and J. Zhang, “Thermodynamic optimization of atmospheric distillation unit,” *Comput. Chem. Eng.*, vol. 103, pp. 201–209, 2017, doi: 10.1016/j.compchemeng.2017.03.024.

[5] Z. Zhang, Z. Wu, D. Rincon, and P. D. Christofides, “Real-time optimization and control of nonlinear processes using machine learning,” *Mathematics*, vol. 7, no. 10, pp. 1–25, 2019, doi: 10.3390/math7100890.

[6] Y. Al-Dunainawi and M. F. Abbod, “PSO-PD fuzzy control of distillation column,” *IntelliSys 2015 - Proc. 2015 SAI Intell. Syst. Conf.*, pp. 554–558, 2015, doi: 10.1109/IntelliSys.2015.7361194.

[7] B. Mamandipoor, M. Majd, S. Sheikhalishahi, C. Modena, and V. Osmani, “Monitoring and detecting faults in wastewater treatment plants using deep learning,” *Environ. Monit. Assess.*, vol. 192, no. 2, pp. 1–12, 2020, doi: 10.1007/s10661-020-8064-1.

[8] C. Li, D. Zhao, S. Mu, W. Zhang, N. Shi, and L. Li, “Fault diagnosis for distillation process based on CNN–DAE,” *Chinese J. Chem. Eng.*, vol. 27, no. 3, pp. 598–604, 2019, doi: 10.1016/j.cjche.2018.12.021.

[9] K. McBride, E. I. Sanchez Medina, and K. Sundmacher, “Hybrid Semi-parametric Modeling in Separation Processes: A Review,” *Chemie-Ingenieur-Technik*, vol. 92, no. 7, pp. 842–855, 2020, doi: 10.1002/cite.202000025.

[10] E. Napolitano, L. Campanile, A. Portolano, L. P. Di Bonito, and F. Di Natale, “Analysis of a marine scrubber operation with conventional and AI based methodologies,” no. September, pp. 1–6, 2022.