Second International Conference Math 2 Product (M2P 2025)
Emerging Technologies in Computational Science
for Industry, Sustainability and Innovation
June 4th -6yh, 2025, Valencia, Spain

APPLICATIONS OF MACHINE LEARNING TO CHEMICAL PROCESSES

F. CALABRÒ[*], D. E. DE FALCO[†], C. EHM[+], R. CIPULLO[+] AND V. BUSICO[+]

[*] Università degli studi di Napoli Federico II,
Dipartimento di Matematica e Applicazioni "Renato Caccioppoli", Napoli, Italy
calabro@unina.it

[†] Scuola Superiore Meridionale, Napoli, Italy
de.defalco@ssmeridionale.it

[+]Università degli studi di Napoli Federico II
Dipartimento di Scienze Chimiche, Napoli, Italy
christian.ehm@unina.it, rcipullo@unina.it, busico@unina.it@unina.it

ABSTRACT

The applications of Machine Learning (ML) to the chemical sciences are a rapidly growing field in computational science, supported by increasing computing power and abundance of data coming from High Throughput Experiments. The results already shown and discussed in the literature during the past decade has clearly given an idea of the potential that these techniques have to solve many theoretical and practical challenges. A perfect example is Catalysts Design, wherein experts attempt to qualitatively identify correlations between a catalyst's structure and its efficiency. This approach has several problems, including the lack of mechanistic understanding of the details of a transformation, and the fact that human cognitive abilities are limited and finding patterns in large collections of data is practically impossible. Catalyst structures can be characterized by descriptors (numerical representations of molecular properties derived from the 3D molecular structure) that quantify the steric and electronic properties of thousands of candidate molecules, and the performances of a given catalyst candidate can be predicted quantitatively by comparing its properties with a computationally derived model trained on experimental data. More in general, in chemical processes the challenge is the joint analysis of data coming from different sources, such as descriptors, temperature measurements taken during experiments and information on the outcome of such experiments. Scientific ML gives the roadmap to include prior knowledge of the problem at hand so to reduce the amount of data required for the training of a given model in supervised learning tasks, and usually results in better generalization capabilities. By the other side, unsupervised learning models such as Auto-Encoders and regression, play an important role in features extraction and can enhance the imposition of more complex physical constraints through automatic generation of new features and loss terms. In this proposed Section, we invite research that focuses on fundamental computational methodologies, their applications, and the collection and analysis of chemical models and databases. We also encourage submissions from researchers working in industrial settings to share the challenges they face and the results they have achieved in practice.