Exploration of Catalytic Pyrolysis with Active Learning

Yannick Ureel[1], Maarten R. Dobbelaere[1], Oğuzhan Akin[1],
Robin John Varghese[1], Kevin M. Van Geem[1]

[1]*Laboratory for Chemical Technology, Ghent University, Belgium*

Experiments are at the basis of scientific research. This is also true for chemical engineering, whether it is for the investigation of process conditions, catalysts or reactors, experiments are required for a thorough scientific understanding. These experiments are often expensive, time-consuming and labor-intensive. Furthermore with the rise of machine learning in chemical engineering accurate experimental data is becoming even more important. Design of experiments (DoE) aims to increase the efficiency of scientific research by extracting maximal information from a minimum of experiments. The combination of DoE with machine learning leads to the field of active learning, which results in a more flexible, multidimensional selection of experiments than regular DoE techniques. Active learning can be employed for the modelling of reactions with well-selected experiments, nevertheless its use is not widespread in chemical engineering [1]. To the best of our knowledge, active learning has not yet been applied in reaction modelling as current methods still require an excessive amount of data for the selection of experiments.

In this work a novel data-scarce active learning methodology for the optimal selection of experiments is proposed. This method combines Gaussian processes and clustering techniques for a flexible determination of the required experiments. In contrast to common active learning strategies, this efficient, data-scarce active learning approach already shows it benefits for a small number (10-20) of experiments. This is a significant advantage of this methodology over other active learning approaches as it enables active learning for a real experimental campaign, where only a limited amount of experiments can be executed. The algorithm is validated on a real experimental campaign to study the effect of several process conditions and catalysts on the catalytic pyrolysis of plastic waste. To proof the benefits of active learning, a widely applied regular DoE technique (full-factor design) is compared to our data-scarce active learning approach.

References

[1] D. Reker, “Practical considerations for active machine learning in drug discovery,” *Drug Discov. Today Technol.*, 2020, doi: 10.1016/j.ddtec.2020.06.001.