



# Modelling of chemical processes by semi-supervised learning

## [MCP-SSL]

### Project Initiator(s)

KU Leuven - prof. dr. Mumin Enis Leblebici - MuminEnis.Leblebici@kuleuven.be

### Project context

[Describe briefly the project (object, cause and effect) in max. 5 lines]

This project is to develop models of chemical processes via semi-supervised learning technology. A robust model is critical for control and optimization. However, the reality always differs from first principle models, and not all measured data is well-explained, especially for spectra analytical tools. With semi-supervised learning technologies, model training and validation use not only ready-identified data but also unidentified data.

### Innovation goal

[Describe briefly the project goal, the novelty & innovative step of the proposed solution (compared to state-of-the-art) in max. 10 lines]

The goal is to improve chemical processes' performance with the existing historical data, which at present is not or only partially used. The current development and validation of the process model usually rely on identified data. However, the identification of every participating data is expensive and difficult. To overcome these challenges, semi-supervised learning technologies are envisioned utilizing a smaller amount of identified data than conventional machine learning models. It can expand the meaningful identifications from identified data into unidentified data (e.g., unknown concentrations), reducing the requirements for the costly identification of the entire data. Semi-supervised learning has been implemented in many classification problems, while it is relatively few in regression tasks. In this way, a low-cost and high-accuracy approach can be established to develop process models efficiently using unidentified data.

### Requested expertise

[Describe briefly your ideal project partner or the missing expertise you are looking for in max. 10 lines]

Chemical or fine chemical companies who would like to improve their current processes owning a large amount of data, which is deemed too expensive or impossible to valorize in every output.

Deployment partners to help deploy the developed models and control agents to company processes.