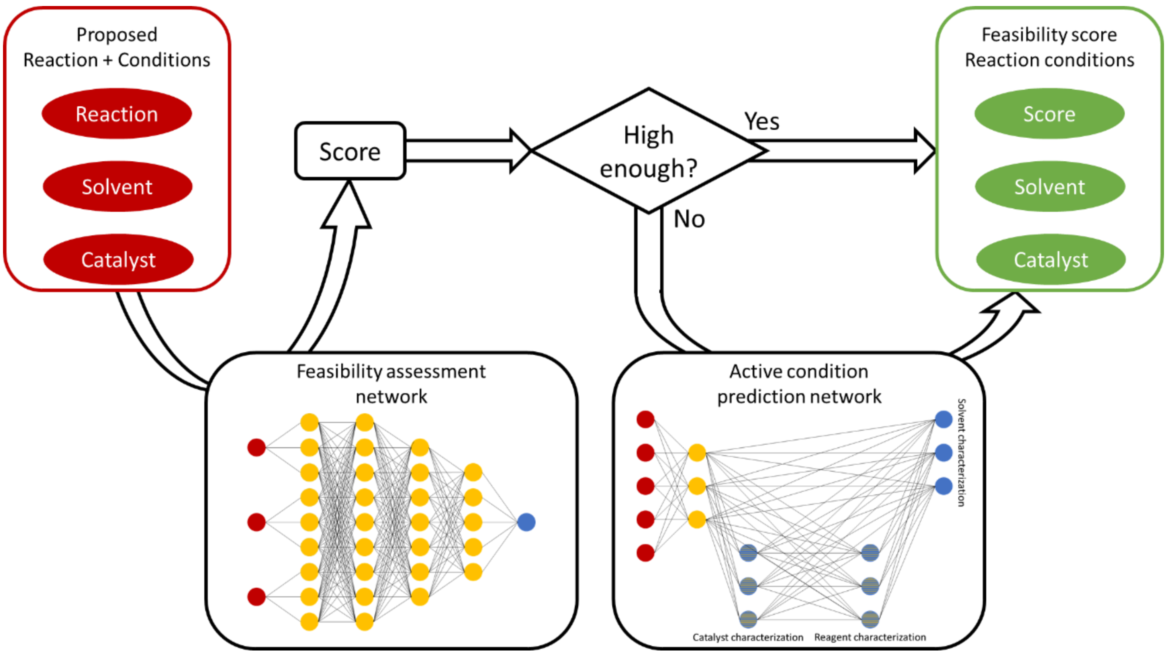
Machine Learning and Continuous Flow: Detection and Correction of Flow-Incompatible Reaction Conditions

Pieter P. Plehiers[1], Connor W. Coley[2], Hanyu Gao[2], William H. Green[2], Guy B. Marin[1], Christian V. Stevens[3] and Kevin M. Van Geem[1,*]

[1]Laboratory for Chemical Technology, Ghent University, Technologiepark 914 9052 Gent, Belgium
[2]Dept. of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139
[3]SynBioC Research Group, Department of Sustainable Organic Chemistry and Technology, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Gent, Belgium

Keywords: *Retrosynthesis, Machine learning, Continuous flow, Prediction, Reaction conditions*

The road to a fully automated robo-chemist is littered with obstacles [1]. One is the synthesis method itself. There is an increasing trend toward using flow chemistry in pharmaceutical synthesis due to the increased safety, operational control and ease of scalability it offers. Despite its advances, performing syntheses in flow demands a more extensive process development to ensure successful operation [2]. Two major issues that must be considered are precipitation causing blockage and high-viscosity flow. Existing retro-synthesis software already takes into account several factors when assessing a synthetic route, such as the likelihood that the desired products are the real products for a proposed reaction [3]. In this work, a method based on artificial neural networks has been developed to detect flow-incompatible reaction conditions, and if necessary, propose the set of conditions that results in the highest possible flow compatibility for that reaction. The detection step consists of using a filtered dataset of reactions and reaction conditions from the Reaxys® database, to train a neural network to determine a compatibility score for the combination of a given reaction and a given set of reaction conditions. If that score is below a specified threshold, a second neural network has been taught to propose those reaction conditions that have the highest likelihood of success in continuous flow. The calculated score can be used to direct the search for synthetic pathways towards the most feasible ones in continuous synthesis. For those reactions for which new conditions are proposed, comparing new and old reaction conditions provides insight into why the initial conditions were considered unfeasible.



Acknowledgements

P.P.P acknowledges financial support from a doctoral fellowship of the Research Foundation – Flanders (FWO). C.W.C received funding from the NSF Graduate Research Fellowship Program under grant no. 1122374.

1. Peplow, M., *Nature* **2014,** 512, (7512), 20.
2. Plutschack, M. B., et al., *Chemical Reviews* **2017,** 117, (18), 11796-11893.
3. Coley, C. W., et al., *ACS Central Science* **2017,** 3, (5), 434-443.

*Corresponding author: Kevin.VanGeem@UGent.be