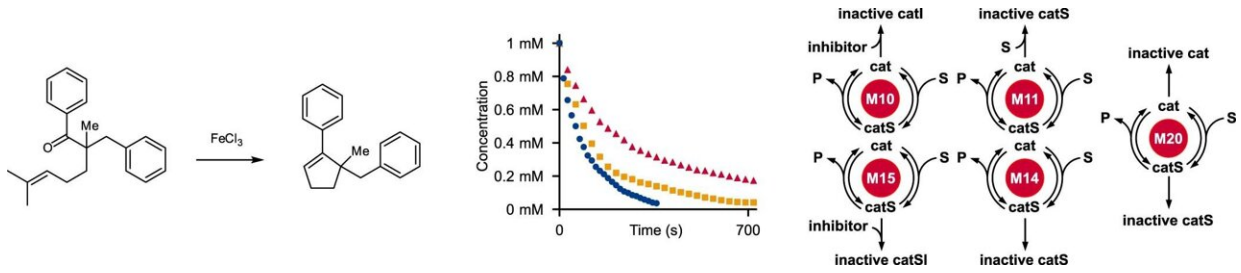
A machine-learning tool that classifies catalytic reactions based on simulated kinetic signatures

February 2 2023, by Bob Yirka



Additional case study with experimental kinetic data. Includes the reaction under study, the experimental kinetic data used as input for the AI-model and its output. Symbols correspond to substrate concentration. Red triangles: lowest catalyst loading; yellow squares: medium catalyst loading; blue circles: largest catalyst loading. Credit: *Nature* (2023). DOI: 10.1038/s41586-022-05639-4

A pair of chemists at the University of Manchester has developed a machine-learning tool that can be used to classify reactions based on simulated kinetic signatures of reactions. In their paper published in the journal *Nature*, Jordi Burés and Igor Larrosa describe combining two deep-learning algorithms to create a system that could speed up the process of new design reactions.

Danilo Lustosa and Anat Milo with Ben-Gurion University of the Negev have published a News and Views piece in the same journal issue

outlining the intricacies involved in building AI applications for chemistry research and discuss the work by the team in the U.K. on this new effort.

To create their AI tool, Burés and Igor Larrosa used two different kinds of neural networks. The first was used to track concentration changes in a reaction over time. The second was used to process data from the first network. The resulting tool, which has not yet been named, was trained using 5 million simulated kinetic samples and their associations with 20 catalytic reaction mechanisms. When running, it is capable of accepting data from an assortment of over a half-million parameters. Once trained, the system applies rules it has learned from the training as a means of classification.

The researchers then tested their model using simulated data and found it produced just 38 errors out of 100,000 samples. Next, they added noise to make the testing more closely resemble real-life scenarios and found the error rate came down to just 99.6% under average conditions and 83% under extreme conditions. The researchers then tested their model using data from prior experiments and found the results to be what they describe as "chemically sound."

Lustosa and Milo note that the tool created by Burés and Larrosa should be able to reduce the bottlenecks typically associated with the production of huge amounts of experimental kinetic data that are generally needed for deep-learning AI systems. They also note that the data produced by the system could likely be used in training efforts, because it is much cleaner than usual due to associations with mechanistic scenarios.

More information: Jordi Burés et al, Organic reaction mechanism classification using machine learning, *Nature* (2023). DOI: 10.1038/s41586-022-05639-4

Danilo M. Lustosa et al, Machine learning classifies catalytic-reaction mechanisms, *Nature* (2023). DOI: 10.1038/d41586-023-00145-7