

# Machine learning model helps characterize compounds for drug discovery

October 14 2020

---



Credit: CC0 Public Domain

Tandem mass spectrometry is a powerful analytical tool used to characterize complex mixtures in drug discovery and other fields.

Now, Purdue University innovators have created a new method of applying machine learning concepts to the tandem [mass spectrometry](#) process to improve the flow of information in the development of new drugs. Their work is published in *Chemical Science*.

"Mass spectrometry plays an integral role in [drug discovery](#) and development," said Gaurav Chopra, an assistant professor of analytical and physical chemistry in Purdue's College of Science. "The specific implementation of bootstrapped machine learning with a small amount of positive and negative training data presented here will pave the way for becoming mainstream in day-to-day activities of automating characterization of compounds by chemists."

Chopra said there are two major problems in the field of machine learning used for chemical sciences. Methods used do not provide chemical understanding of the decisions that are made by the algorithm, and new methods are not typically used to do blind experimental tests to see if the proposed models are accurate for use in a chemical laboratory.

"We have addressed both of these items for a methodology that is isomer selective and extremely useful in [chemical sciences](#) to characterize complex mixtures, identify [chemical reactions](#) and drug metabolites, and in fields such as proteomics and metabolomics," Chopra said.

The Purdue researchers created statistically robust machine learning models to work with less training data—a technique that will be useful for drug discovery. The model looks at a common neutral reagent—called 2-methoxypropene (MOP) - and predicts how compounds will interact with MOP in a tandem mass spectrometer in order to obtain structural information for the compounds.

"This is the first time that machine learning has been coupled with diagnostic gas-phase ion-molecule reactions, and it is a very powerful

combination, leading the way to completely automated mass spectrometric identification of organic compounds," said Hilkka Kenttämä, the Frank Brown Distinguished Professor of Analytical Chemistry and Organic Chemistry. "We are now introducing many new reagents into this method."

The Purdue team introduces chemical reactivity flowcharts to facilitate [chemical](#) interpretation of the decisions made by the [machine learning](#) method that will be useful to understand and interpret the mass spectra for structural information.

**More information:** Jonathan Fine et al, Graph-based machine learning interprets and predicts diagnostic isomer-selective ion–molecule reactions in tandem mass spectrometry, *Chemical Science* (2020). [DOI: 10.1039/D0SC02530E](#)

Provided by Purdue University

Citation: Machine learning model helps characterize compounds for drug discovery (2020, October 14) retrieved 23 April 2024 from <https://phys.org/news/2020-10-machine-characterize-compounds-drug-discovery.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.