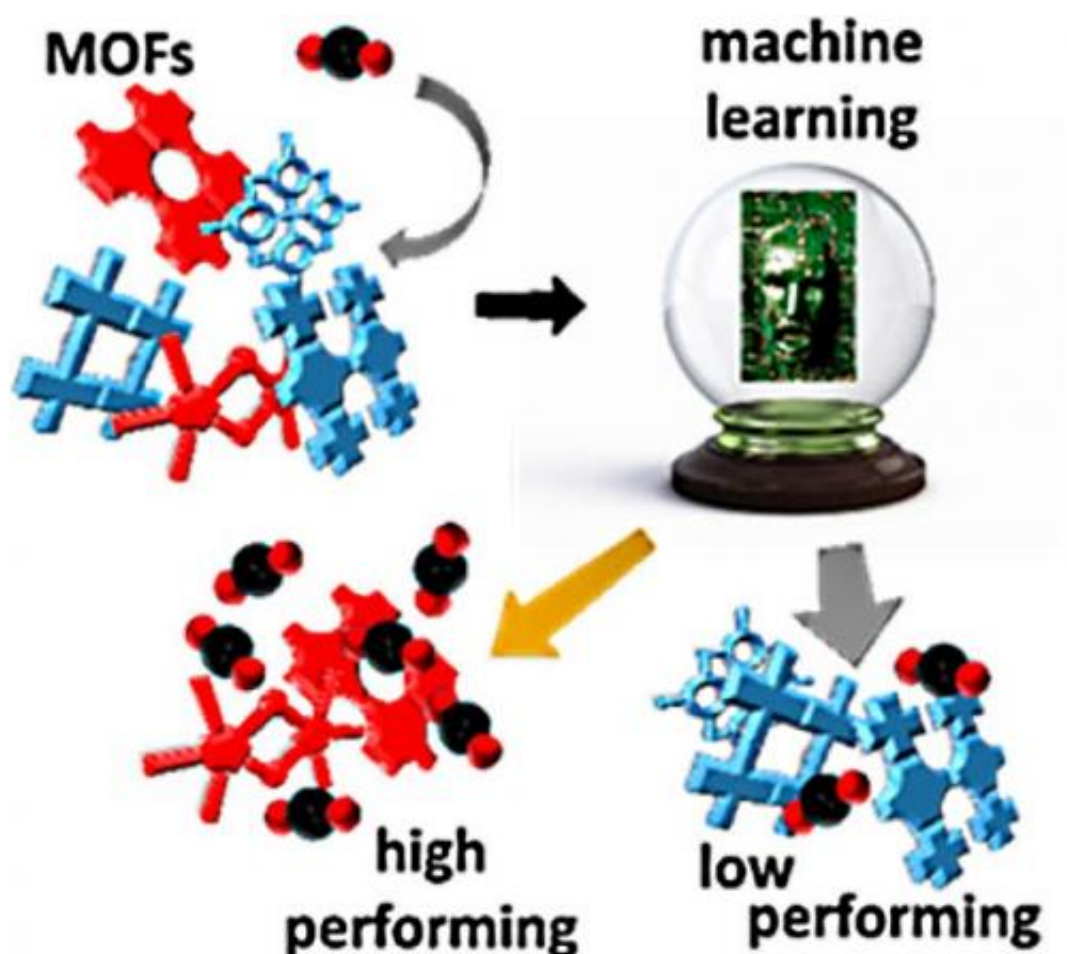


# Machine learning algorithm makes impossible screening of advanced materials possible

September 15 2014, by Lisa Zyga



Using machine learning, a new model can screen very large numbers of different metal-organic frameworks in a fraction of the time it currently takes using compute-intensive methods. Credit: Fernandez, et al. ©2014 American Chemical Society

(Phys.org) —A fundamental part of climate change response is expected to involve the discovery of advanced materials capable of cost-effectively capturing CO<sub>2</sub> from burning fossil fuels. One particular class of materials, called metal-organic frameworks (MOFs), have tremendous potential to revolutionize CO<sub>2</sub> capture technologies. For example, MOFs have record-breaking high surface areas, so that one gram of solid can have two football fields of internal surface area to adsorb and later release CO<sub>2</sub>.

MOFs are formed from the building blocks made of metal ions and organic linkers. Because of the many types of [building blocks](#) that can be combined in different ways, there are literally billions of hypothetical MOF structures. Despite the fact that advanced computational simulations can accurately calculate the CO<sub>2</sub> adsorption capacity of hypothetical MOFs, wading through all possible structures using compute-intensive simulations in order to identify the best candidate materials would take far too much time to make exhaustive screening practical.

To overcome this computing limitation, researchers at the University of Ottawa in Ottawa, Ontario, have used [machine learning](#) tools to develop models that can rapidly identify the best performing MOFs for CO<sub>2</sub> capture in a fraction of the time it would take to screen the entire database. By pre-screening for the top MOFs, the model greatly reduces the number of MOFs that require more intensive screening, and could decrease the overall computing time by an order of magnitude. The study is published in a recent issue of *The Journal of Physical Chemistry Letters*.

"The greatest significance of the work is that we have developed a way to enable huge libraries of candidate materials to be screened in a practical amount of time," Tom Woo, Professor at the University of Ottawa, told *Phys.org*. "This may eventually lead to the development of

new materials that will make scrubbing CO<sub>2</sub> from burning coal a practical and cost effective endeavor."

In the study, the researchers generated a database of 320,000 MOFs by combining various [metal ions](#) and organic linkers. Then they randomly selected 10% of the database to form the calibration set to train the new computer model using machine learning algorithms. Using the structure of the material, the model learned to classify MOFs as "high-performing" if the predicted CO<sub>2</sub> uptake capacity exceeded a predetermined threshold value, or as "low-performing" if it did not. The machine learning models could classify a MOF in a fraction of a second whereas the compute intensive simulations used to predict the CO<sub>2</sub> uptake could take hours to complete.

After the training process, the researchers tested the model on the remaining 290,000 MOFs of the database. All of these MOFs had previously been evaluated for CO<sub>2</sub> uptake capacity by the compute-intensive simulations, allowing for a detailed analysis of the model's accuracy. The model was able to recover 945 of the top 1000 high performing MOFs while only flagging 10% of the database for more compute intensive screening.

Translating these recovery rates to a more concrete example, the researchers looked at the case of a database of 1.5 million MOFs generated from a relatively small set of components. Using the model to prescreen the database would flag just 150,000 structures while recovering about 95% of the actual high-performing candidates. This would reduce number of compute intensive simulations performed from 1.5 million to 150,000, thereby making a screening of the library feasible.

The approach could be applied more generally to large libraries of MOF structures for other applications, as long as data is available to calibrate

the machine learning models.

"We are developing other machine learning models to screen for other gas separations important for clean energy applications, and we are using the models to screen the massive databases of materials we have created," Woo said.

**More information:** Michael Fernandez, et al. "Rapid and Accurate Machine Learning Recognition of High Performing Metal Organic Frameworks for CO<sub>2</sub> Capture." *The Journal of Physical Chemistry Letters*. DOI: [10.1021/jz501331m](https://doi.org/10.1021/jz501331m)

© 2014 Phys.org

Citation: Machine learning algorithm makes impossible screening of advanced materials possible (2014, September 15) retrieved 3 May 2024 from <https://phys.org/news/2014-09-machine-algorithm-impossible-screening-advanced.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.
---