Machine learning boosts the search for 'superhard' materials

December 17 2020, by Jeannie Kever

Researchers have developed a machine learning model that can accurately predict the hardness of new materials, allowing scientists to more readily find compounds suitable for use in a variety of applications. Credit: University of Houston

Superhard materials are in high demand in industry, from energy production to aerospace, but finding suitable new materials has largely been a matter of trial and error based on classical materials such as diamonds. Until now.

Researchers from the University of Houston and Manhattan College have reported a [machine learning model](#) that can accurately predict the hardness of new materials, allowing scientists to more readily find compounds suitable for use in a variety of applications. The work was reported in *Advanced Materials*.

Materials that are superhard—defined as those with a hardness value exceeding 40 gigapascals on the Vickers scale, meaning it would take more than 40 gigapascals of pressure to leave an indentation on the material's surface—are rare.

"That makes identifying new materials challenging," said Jakoah Brgoch, associate professor of chemistry at UH and corresponding author for the paper. "That is why materials like synthetic diamond are still used even though they are challenging and expensive to make."

One of the complicating factors is that the hardness of a material may vary depending on the amount of pressure exerted, known as load dependence. That makes testing a material experimentally complex and using computational modeling today almost impossible.

The model reported by the researchers overcomes that by predicting the load-dependent Vickers hardness based solely on the chemical composition of the material. The researchers report finding more than 10 new and promising stable borocarbide phases; work is now underway to design and produce the materials so they can be tested in the lab.

Based on the model's reported accuracy, the odds are good. Researchers reported the accuracy at 97%.

First author Ziyan Zhang, a doctoral student at UH, said the database built to train the algorithm is based on data involving 560 different compounds, each yielding several data points. Finding the data required poring over hundreds of published academic papers to find data needed to build a representative dataset.

"All good machine learning projects start with a good dataset," said Brgoch, who is also a principal investigator with the Texas Center for Superconductivity at UH. "The true success is largely the development of this dataset."

In addition to Brgoch and Zhang, additional researchers on the project include Aria Mansouri Tehrani and Blake Day, both with UH, and Anton O. Oliynyk from Manhattan College.

Researchers traditionally have used machine learning to predict a single variable of hardness, Brgoch said, but that doesn't account for the complexities of the property like load dependence, which he said still aren't well understood. That makes machine learning a good tool, despite earlier limitations.

"A machine learning system doesn't need to understand the physics," he said. "It just analyzes the training data and makes new predictions based on statistics."

Machine learning does have limitations, though.

"The idea of using machine learning isn't to say, 'Here is the next greatest material,' but to help guide our experimental search," Brgoch said. "It tells you where you should look."

More information: Ziyan Zhang et al. Finding the Next Superhard Material through Ensemble Learning, *Advanced Materials* (2020). DOI: 10.1002/adma.202005112

Provided by University of Houston