Machine learning for solar energy is a supercomputer killer

June 23 2021

Credit: CC0 Public Domain

Supercomputers could find themselves out of a job thanks to a suite of

new machine learning models that produce rapid, accurate results using a normal laptop.

Researchers at the ARC Centre of Excellence in Exciton Science, based at RMIT University, have written a [program](#) that predicts the band gap of materials, including for solar energy applications, via freely available and easy-to-use software. Band gap is a crucial indication of how efficient a material will be when designing new solar cells.

Band gap predictions involve quantum and atomic-scale chemical calculations and are often made using density functional theory. Until now, this process has required hundreds of hours of costly [supercomputer](#) processing time, as well as complicated and expensive software.

To address this issue, the researchers trained a [machine learning model](#) using data generated from 250,000 previous supercomputer calculations. The results have been published in *Journal of Cheminformatics*.

Significantly, while the program is capable of including multiple variables, it was found that just one factor, stoichiometry, contains—in almost all cases—enough information to accurately predict band gap. Stoichiometry is the numerical relationships between chemical reactants and products, like the volume of ingredients in a recipe to bake a cake.

More work is needed to fully understand why stoichiometry alone proved to be so useful. But it raises the exciting prospect of lengthy supercomputer calculations no longer being required for some applications. The [artificial neural network](#) that powers the machine learning programs could one day be succeeded by a [software program](#) that performs a similar function to density functional theory, albeit with far more simplicity.

Lead author Carl Belle says that "if you want to do simulations but you need to have millions of dollars of supercomputing infrastructure behind you, you can't do it. If we can dig into why the stoichiometric configuration is so powerful, then it could mean that supercomputers are not needed to screen candidate materials, nor for accurate simulations. It could really open things up to a whole new group of scientists to use."

The machine learning program isn't limited to band gap. It can be used to predict the properties of many other materials for other contexts, and has been developed by a professional programmer, making it useful not only for scientists and academics but also for businesses and corporate applications.

"It's built to industry standard and it's designed to be collaborative," Belle said.

"The website has a fully relational database. It's got millions of records. It's all there and freely available to use. We're ready to go."