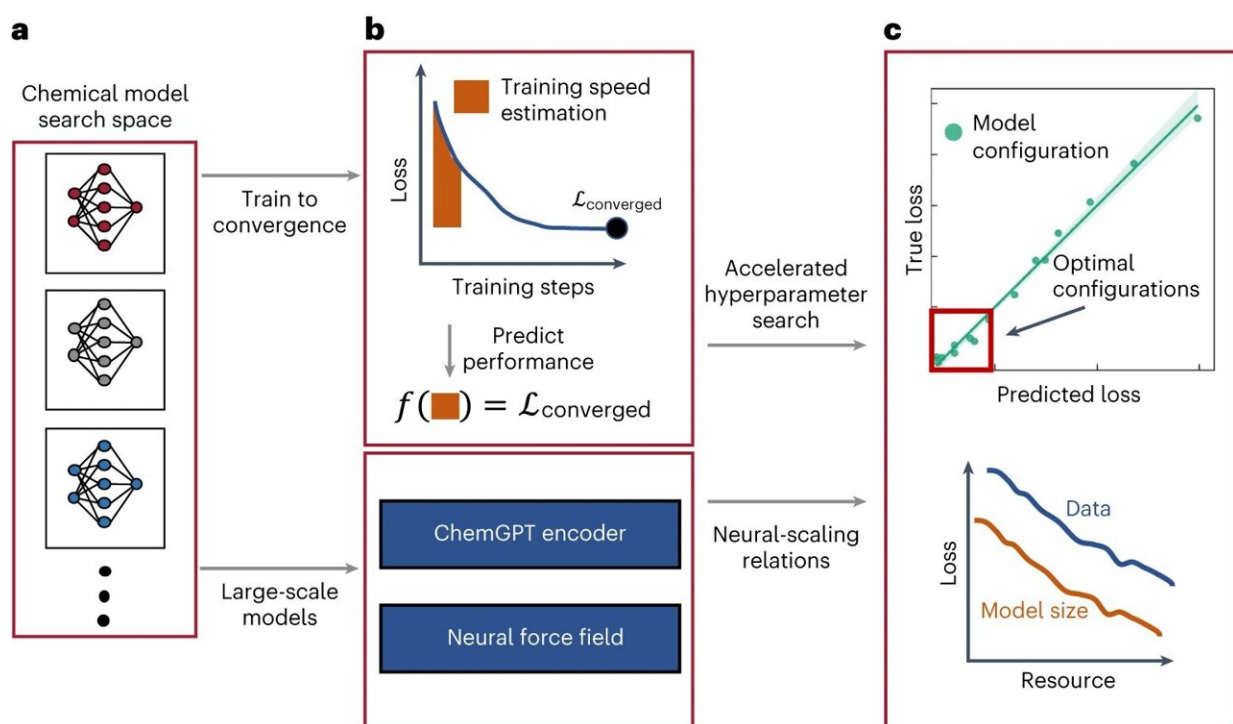


Study explores the scaling of deep learning models for chemistry research

November 11 2023, by Ingrid Fadelli



Discovery of neural-scaling relations for deep chemical models. **a,b**, Over a domain of model candidates (**a**), final, converged model loss is predicted from only a few initial epochs of training for large-scale models (**b**). **c**, Non-optimal model architectures and hyperparameter configurations are identified early in training, allowing for efficient selection of the ideal architecture and hyperparameters. The model with the best hyperparameters is then trained with varying model and dataset sizes to discover neural-scaling relations. Credit: *Nature Machine Intelligence* (2023). DOI: 10.1038/s42256-023-00740-3

Deep neural networks (DNNs) have proved to be highly promising tools for analyzing large amounts of data, which could speed up research in various scientific fields. For instance, over the past few years, some computer scientists have trained models based on these networks to analyze chemical data and identify promising chemicals for various applications.

Researchers at the Massachusetts Institute of Technology (MIT) recently carried out a study investigating the neural scaling behavior of large DNN-based models trained to generate advantageous [chemical](#) compositions and learn interatomic potentials. Their paper, [published](#) in *Nature Machine Intelligence*, shows how quickly the performance of these models can improve as their size and the pool of data they are trained on are increased.

"The paper '[Scaling Laws for Neural Language Models](#)' by Kaplan et al., was the main inspiration for our research," Nathan Frey, one of the researchers who carried out the study, told Tech Xplore. "That paper showed that increasing the size of a [neural network](#) and the amount of data it's trained on leads to predictable improvements in model training. We wanted to see how 'neural scaling' applies to models trained on chemistry data, for applications like drug discovery."

Frey and his colleagues started working on this research project back in 2021, thus before the release of the renowned AI-based platforms ChatGPT and Dall-E 2. At the time, the future upscaling of DNNs was perceived as particularly relevant to some fields and studies exploring their scaling in the physical or life sciences were scarce.

The researchers' study explores the neural scaling of two distinct types of models for chemical data analysis: a large language model (LLM) and a graph neural network (GNN)-based model. These two different types of models can be used to generate chemical compositions and learn the

potentials between different atoms in chemical substances, respectively.

"We studied two very different types of models: an autoregressive, GPT-style language model we built called 'ChemGPT' and a family of GNNs," Frey explained. "ChemGPT was trained in the same way ChatGPT is, but in our case ChemGPT is trying to predict the next token in a string that represents a molecule. The GNNs are trained to predict the energy and forces of a molecule."

To explore the scalability of the ChemGPT model and of GNNs, Frey and his colleagues explored the effects of a model's size and the size of the dataset used to train it on various relevant metrics. This allowed them to derive a rate at which these models improve as they become larger and are fed more data.

"We do find 'neural scaling behavior' for chemical models, reminiscent of the scaling behavior seen in LLM and vision models for various applications," Frey said.

"We also showed that we are not near any kind of fundamental limit for scaling chemical models, so there is still a lot of room to investigate further with more compute and bigger datasets. Incorporating physics into GNNs via a property called 'equivariance' has a dramatic effect on improving scaling efficiency, which is an exciting result because it's actually quite difficult to find algorithms that change scaling behavior."

Overall, the findings gathered by this team of researchers shed new light on the potential of two types of AI models for conducting chemistry research, showing the extent to which their performance can improve as they are scaled up. This work could soon inform additional studies exploring the promise and margin for improvement of these models, as well as that of other DNN-based techniques for specific scientific applications.

"Since our work first appeared, there has already been exciting follow up work probing the capabilities and limitations of scaling for chemical models," Frey added. "More recently, I have also been working on [generative models for protein design](#) and thinking about how scaling impacts models for biological data."

More information: Nathan C. Frey et al, Neural scaling of deep chemical models, *Nature Machine Intelligence* (2023). [DOI: 10.1038/s42256-023-00740-3](#)

© 2023 Science X Network

Citation: Study explores the scaling of deep learning models for chemistry research (2023, November 11) retrieved 29 April 2024 from <https://techxplore.com/news/2023-11-explores-scaling-deep-chemistry.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.