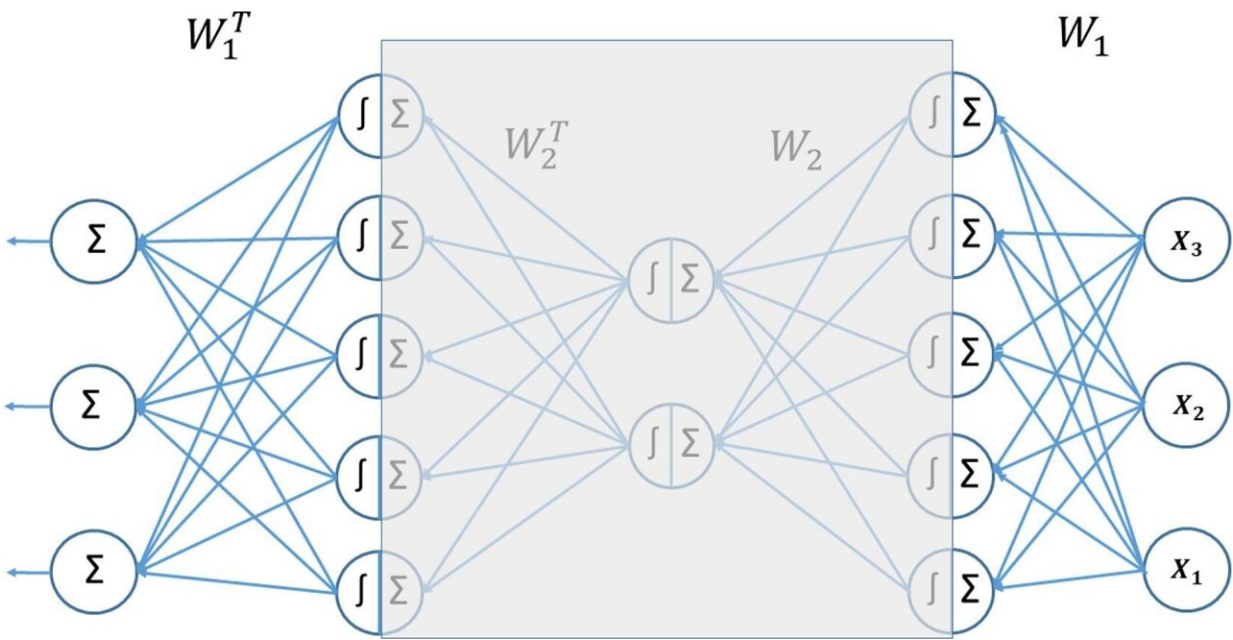


New technique based on 18th-century mathematics shows simpler AI models don't need deep learning

October 5 2023



Layerwise pretraining from heads to inner layers. The most outer layer is trained first and its residual is then fed as training data for the next hidden layer until all layers have been sequentially pretrained. Credit: *Neurocomputing* (2023). DOI: 10.1016/j.neucom.2023.126520

Researchers from the University of Jyväskylä were able to simplify the most popular technique of artificial intelligence, deep learning, using 18th-century mathematics. They also found that classical training

algorithms that date back 50 years work better than the more recently popular techniques. Their simpler approach advances green IT and is easier to use and understand.

The recent success of artificial intelligence is significantly based on the use of one core technique: [deep learning](#). Deep learning refers to [artificial intelligence](#) techniques where networks with a large number of data processing layers are trained using massive datasets and a substantial amount of computational resources.

Deep learning enables computers to perform [complex tasks](#) such as analyzing and generating images and music, playing digitized games and, most recently in connection with ChatGPT and other generative AI techniques, acting as a [natural language](#) conversational agent that provides high-quality summaries of existing knowledge.

Six years ago, Professor Tommi Kärkkäinen and Doctoral researcher Jan Hänninen conducted preliminary studies on data reduction. The results were surprising: If one combines simple network structures in a novel way then depth is not needed. Similar or even better results can be obtained with shallow models.

"The use of deep learning techniques is a complex and error-prone endeavor, and the resulting models are difficult to maintain and interpret," says Kärkkäinen. "Our new model in its shallow form is more expressive and can reliably reduce large datasets while maintaining all the necessary information in them."

The structure of the new AI technique dates back to 18th-century mathematics. Kärkkäinen and Hänninen also found that the traditional optimization methods from the 1970s work better in preparing their model compared to the 21st-century techniques used in deep learning.

"Our results ensure that the use of neural networks in various applications is easier and more reliable than before," suggests Hänninen. [The study](#) is published in the journal *Neurocomputing*.

Simpler models lead to greener and more ethical AI

Artificial intelligence has an increasingly significant role in modern technologies and, therefore, it is more and more important to understand how AI does what it does.

"The more transparent and simpler AI is, the easier it is to consider its ethical use," says Kärkkäinen. "For instance, in [medical applications](#) deep learning techniques are so complex that their straightforward use can jeopardize patient safety due to unexpected, hidden behavior."

The researchers note that simpler models can help develop green IT and are more environmentally friendly because they save computational resources and use significantly less energy.

The results, which challenge the common beliefs and currently popular perceptions with deep learning techniques, were difficult to get published.

"Deep learning has such a prominent role in research, development, and the AI business that, even if science always progresses and reflects the latest evidence, the community itself may have resistance to change."

"We are very interested to see how these results will be received in the scientific and business community," says Kärkkäinen. "Our new AI has a range of applications in our own research, from nanotechnology for better materials in the sustainable economy to improving digital learning environments and increasing the reliability and transparency of medical and well-being technology."

More information: Tommi Kärkkäinen et al, Additive autoencoder for dimension estimation, *Neurocomputing* (2023). [DOI: 10.1016/j.neucom.2023.126520](https://doi.org/10.1016/j.neucom.2023.126520)

Provided by University of Jyväskylä

Citation: New technique based on 18th-century mathematics shows simpler AI models don't need deep learning (2023, October 5) retrieved 29 April 2024 from <https://techxplore.com/news/2023-10-technique-based-18th-century-mathematics-simpler.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.