# AI has a large and growing carbon footprint, but there are potential solutions on the horizon

February 19 2024, by Shirin Dora

Credit: AI-generated image

Given the huge problem-solving potential of artificial intelligence (AI), it wouldn't be far-fetched to think that AI could also help us in tackling the climate crisis. However, when we consider the energy needs of AI models, it becomes clear that the technology is as much a part of the climate problem as a solution.

The emissions come from the infrastructure associated with AI, such as building and running the data centers that handle the large amounts of information required to sustain these systems.

But different technological approaches to how we build AI systems could help reduce its carbon footprint. Two technologies in particular hold promise for doing this: spiking neural networks and lifelong learning.

The lifetime of an AI system can be split into two phases: training and inference. During training, a relevant dataset is used to build and tune—improve—the system. In inference, the trained system generates predictions on previously unseen data.

For example, training an AI that's to be used in self-driving cars would require a dataset of many different driving scenarios and decisions taken by human drivers.

After the training phase, the AI system will predict effective maneuvers for a self-driving car. Artificial neural networks (ANN), are an underlying technology used in most current AI systems.

They have many different elements to them, called parameters, whose values are adjusted during the training phase of the AI system. These parameters can run to more than 100 billion in total.

While large numbers of parameters improve the capabilities of ANNs, they also make training and inference resource-intensive processes. To put things in perspective, training GPT-3 (the precursor AI system to the current ChatGPT) generated 502 metric tons of carbon, which is equivalent to driving 112 petrol powered cars for a year.

GPT-3 further emits 8.4 tons of $CO_2$ annually due to inference. Since the AI boom started in the early 2010s, the energy requirements of AI systems known as large language models (LLMs)—the type of technology that's behind ChatGPT—have gone up by a factor of 300,000
.

With the increasing ubiquity and complexity of AI models, this trend is going to continue, potentially making AI a significant contributor of $CO_2$ emissions. In fact, our current estimates could be lower than AI's actual carbon footprint due to a lack of standard and accurate techniques for measuring AI-related emissions.

## Spiking neural networks

The previously mentioned new technologies, spiking neural networks (SNNs) and lifelong learning (L2), have the potential to lower AI's ever-increasing carbon footprint, with SNNs acting as an energy-efficient alternative to ANNs.

ANNs work by processing and learning patterns from data, enabling them to make predictions. They work with decimal numbers. To make accurate calculations, especially when multiplying numbers with decimal points together, the computer needs to be very precise. It is because of these decimal numbers that ANNs require lots of computing power, memory and time.

This means ANNs become more energy-intensive as the networks get larger and more complex. Both ANNs and SNNs are inspired by the brain, which contains billions of neurons (nerve cells) connected to each other via synapses.

Like the brain, ANNs and SNNs also have components which researchers call neurons, although these are artificial, not biological ones.

The key difference between the two types of neural networks is in the way individual neurons transmit information to each other.

Neurons in the human brain communicate with each other by transmitting intermittent electrical signals called spikes. The spikes themselves do not contain information. Instead, the information lies in the timing of these spikes. This binary, all-or-none characteristic of spikes (usually represented as 0 or 1) implies that neurons are active when they spike and inactive otherwise.

This is one of the reasons for energy efficient processing in the brain.

Just as Morse code uses specific sequences of dots and dashes to convey messages, SNNs use patterns or timings of spikes to process and transmit information. So, while the artificial neurons in ANNs are always active, SNNs consume energy only when a spike occurs.

Otherwise, they have closer to zero energy requirements. SNNs can be up to 280 times more energy efficient than ANNs.

My colleagues and I are developing learning algorithms for SNNs that may bring them even closer to the energy efficiency exhibited by the brain. The lower computational requirements also imply that SNNs might be able to make decisions more quickly.

These properties render SNNs useful for broad range of applications, including space exploration, defense and self-driving cars because of the limited energy sources available in these scenarios.

L2 is another strategy for reducing the overall energy requirements of ANNs over the course of their lifetime that we are also working on.

Training ANNs sequentially (where the systems learn from sequences of data) on new problems causes them to forget their previous knowledge while learning new tasks. ANNs require retraining from scratch when their operating environment changes, further increasing AI-related emissions.

L2 is a collection of algorithms that enable AI models to be trained sequentially on multiple tasks with little or no forgetting. L2 enables models to learn throughout their lifetime by building on their existing knowledge without having to retrain them from scratch.

The field of AI is growing fast and other potential advancements are emerging that can mitigate the energy demands of this technology. For instance, building smaller AI models that exhibit the same predictive capabilities as that of a larger model.

Advances in quantum computing—a different approach to building computers that harnesses phenomena from the world of quantum physics—would also enable faster training and inference using ANNs and SNNs. The superior computing capabilities offered by quantum computing could allow us to find energy-efficient solutions for AI at a much larger scale.

The climate change challenge requires that we try to find solutions for rapidly advancing areas such as AI before their carbon footprint becomes too large.



Provided by The Conversation

Citation: AI has a large and growing carbon footprint, but there are potential solutions on the

horizon (2024, February 19) retrieved 29 April 2024 from
https://techxplore.com/news/2024-02-ai-large-carbon-footprint-potential.html