

# AI's carbon footprint problem

3 July 2020



Credit: Stocksy/Jetta Productions

For all the advances enabled by artificial intelligence, from speech recognition to self-driving cars, AI systems consume a lot of power and can generate high volumes of climate-changing carbon emissions.

A study last year found that training an off-the-shelf AI language-processing system produced 1,400 pounds of emissions—about the amount produced by flying one person roundtrip between New York and San Francisco. The full suite of experiments needed to build and train that AI language system from scratch can generate even more: up to 78,000 pounds, depending on the source of [power](#). That's twice as much as the average American exhales over an entire lifetime.

But there are ways to make machine learning cleaner and greener, a movement that has been called "Green AI." Some algorithms are less power-hungry than others, for example, and many [training sessions](#) can be moved to [remote locations](#) that get most of their power from renewable sources.

The key, however, is for AI developers and companies to know how much their machine learning experiments are spewing and how much those volumes could be reduced.

Now, a team of researchers from Stanford,

Facebook AI Research, and McGill University has come up with an easy-to-use tool that quickly measures both how much electricity a machine learning project will use and how much that means in carbon emissions.

"As machine learning systems become more ubiquitous and more resource intensive, they have the potential to significantly contribute to carbon emissions," says Peter Henderson, a Ph.D. student at Stanford in [computer science](#) and the lead author. "But you can't solve a problem if you can't measure it. Our system can help researchers and industry engineers understand how carbon-efficient their work is, and perhaps prompt ideas about how to reduce their carbon footprint."

## Tracking Emissions

Henderson teamed up on the "experiment impact tracker" with Dan Jurafsky, chair of linguistics and professor of computer science at Stanford; Emma Brunskill, an assistant professor of computer science at Stanford; Jieru Hu, a software engineer at Facebook AI Research; Joelle Pineau, a professor of computer science at McGill and co-managing director of Facebook AI Research; and Joshua Romoff, a Ph.D. candidate at McGill.

"There's a big push to scale up machine learning to solve bigger and bigger problems, using more compute power and more data," says Jurafsky. "As that happens, we have to be mindful of whether the benefits of these heavy-compute models are worth the cost of the impact on the environment."

Machine learning systems build their skills by running millions of statistical experiments around the clock, steadily refining their models to carry out tasks. Those training sessions, which can last weeks or even months, are increasingly power-hungry. And because the costs have plunged for both computing power and massive datasets, machine learning is increasingly pervasive in business, government, academia, and personal life.

To get an accurate measure of what that means for carbon emissions, the researchers began by measuring the power consumption of a particular AI model. That's more complicated than it sounds, because a single machine often trains several models at the same time, so each training session has to be untangled from the others. Each training session also draws power for shared overhead functions, such as data storage and cooling, which need to be properly allocated.

The next step is to translate power consumption into carbon emissions, which depend on the mix of renewable and fossil fuels that produced the electricity. That mix varies widely by location as well as by time of day. In areas with a lot of solar power, for example, the carbon intensity of electricity goes down as the sun climbs higher in the sky.

To get that information, the researchers scoured public sources of data about the energy mix in different regions of the United States and the world. In California, the experiment-tracker plugs into real-time data from California ISO, which manages the flow of electricity over most of the state's grids. At 12:45 p.m. on a day in late May, for example, renewables were supplying 47% of the state's power.

The location of an AI training session can make a big difference in its carbon emissions. The researchers estimated that running a session in Estonia, which relies overwhelmingly on shale oil, will produce 30 times the volume of carbon as the same session would in Quebec, which relies primarily on hydroelectricity.

### Greener AI

Indeed, the researchers' first recommendation for reducing the carbon footprint is to move training sessions to a location supplied mainly by renewable sources. That can be easy, because datasets can be stored on a cloud server and accessed from almost anywhere.

In addition, however, the researchers found that some machine learning algorithms are bigger energy hogs than others. At Stanford, for example,

more than 200 students in a class on reinforcement learning were asked to implement common algorithms for a homework assignment. Though two of the algorithms performed equally well, one used far more power. If all the students had used the more efficient algorithm, the researchers estimated they would have reduced their collective power consumption by 880 kilowatt-hours—about what a typical American household uses in a month.

The result highlights the opportunities for reducing [carbon emissions](#) even when it's not practical to move work to a [carbon](#)-friendly location. That is often the case when machine learning systems are providing services in real time, such as car navigation, because long distances cause communication lags or "latency."

Indeed, the researchers have incorporated an easy-to-use tool into the tracker that generates a website for comparing the energy efficiency of different models. One simple way to conserve energy, they say, would be to establish the most efficient program as the default setting when choosing which one to use.

"Over time," says Henderson, "it's likely that machine learning systems will consume even more energy in production than they do during training. The better that we understand our options, the more we can limit potential impacts to the environment."

The experiment impact tracker is [available online](#) for researchers. It is already being used at the SustainNLP workshop at this year's Conference on Empirical Methods in Natural Language Processing, where researchers are encouraged to build and publish energy-efficient NLP algorithms. The research, which has not been peer-reviewed, was published on the preprint site Arxiv.org.

**More information:** Energy and Policy Considerations for Deep Learning in NLP. [arxiv.org/pdf/1906.02243.pdf](https://arxiv.org/pdf/1906.02243.pdf)

Provided by Stanford University

APA citation: AI's carbon footprint problem (2020, July 3) retrieved 30 September 2022 from <https://techxplore.com/news/2020-07-ai-carbon-footprint-problem.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.*