The environmental impact of AI: a case study

Published 15 June 2023

Christopher Gannatti, CFA
Global Head of Research

In our previous blog, Will AI workloads consume all the world's energy?, we looked at the relationship between increasing processing power and an increase in energy demand, and what this means for artificial intelligence (AI) from an environmental standpoint. In this latest blog, we aim to further illuminate this discussion with a case study of the world's biggest large language model (LLM), BLOOM.

Case study on environmental impact: BLOOM

An accurate estimate of the environmental impact of an LLM being run is far from a simple exercise. One must understand, first, that there is a general 'model life cycle.' Broadly, the model life cycle could be thought of as three phases1:

Inference: This is the phase when a given model is said to be 'up-and-running.' If one is thinking of Google's machine translation system, for example, inference is happening when the system is providing translations for users. The energy usage for any single request is small, but if the overall system is processing 100 billion words per day, the overall energy usage could still be quite large.

Training: This is the phase when the parameters of a model have been set and the system is exposed to data from which it is able to learn such that outputs in the inference phase are judged to be 'accurate'. There are cases where the greenhouse gas emissions impact for training large, cutting-edge models can be comparable to the lifetime emissions of a car.

Model development: This is the phase when developers and researchers are seeking to build the model and will tend to experiment with all sorts of different options. It is easier to measure the impact of training a finished model that becomes public, as opposed to seeking to measure the impact of the research and development process, which might have included many different paths prior to getting to the finished model that the public actually sees.

Therefore, the BLOOM case study focuses on the impact from training the model.

- BLOOM is trained on 1.6 terabytes of data in 46 natural languages and 13 programming languages.
- Note, at the time of the study, Nvidia did not disclose the carbon intensity of this specific chip, so the researchers needed to compile data from a close approximate equivalent setup. It's an important detail to keep in mind, in that an accurate depiction of the carbon impact of training a single model requires a lot of information and, if certain data along the way is not disclosed, there must be more and more estimates and approximations (which will impact the final data).

Figure 1: Summary statistics from training the BLOOM model

Source: Luccioni et al. "Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model." ARXIV.org. Submitted 3 November 2022.

If AI workloads are always increasing, does that mean carbon emissions are also always increasing2?

Considering all data centres, data transmission networks, and connected devices, it is estimated that there were about 700 million tonnes of carbon dioxide equivalent in 2020, roughly 1.4% of global emissions. About two-thirds of the emissions came from operational energy use. Even if 1.4% is not yet a significant number relative to the world's total, growth in this area can be fast.

Currently, it is not possible to know exactly how much of this 700 million tonne total comes directly from AI and machine learning. One possible assumption to make, to come to a figure, is that AI and machine learning workloads were occurring almost entirely in hyperscale data centres. These specific data centres contributed roughly 0.1% to 0.2% of greenhouse gas emissions.

Some of the world's largest firms directly disclose certain statistics to show that they are environmentally conscious. Meta Platforms represents a case in point. If we consider its specific activities:

- Overall data centre energy use was increasing 40% per year from 2016.
- Overall training activity in machine learning was growing roughly 150% per year.
- Overall inference activity was growing 105% per year.
- But Meta Platforms' overall greenhouse gas emissions footprint was down 90% from 2016 due to its renewable energy purchases.

The bottom line is, if companies just increased their compute usage to develop, train and run models—increasing these activities all the time—then it would make sense to surmise that their greenhouse gas emissions would always be rising. However, the world's biggest companies want to be seen as 'environmentally conscious', and they frequently buy renewable energy and even carbon credits. This makes the total picture less clear; whilst there is more AI and it may be more energy intensive in certain respects, if more and more of the energy is coming from renewable sources, then the environmental impact may not increase at anywhere near the same rate.

Conclusion—a fruitful area for ongoing analysis

One of the interesting areas for future analysis will be to gauge the impact of internet search with generative AI versus the current, more standard search process. There are estimates that the carbon footprint of generative AI search could be four or five times higher, but looking solely at this one datapoint could be misleading. For instance, if generative AI search actually saves time or reduces the overall number of searches, in the long run, more efficient generative AI search may help the picture more than it hurts3.

Just as we are currently learning how and where generative AI will help businesses, we are constantly learning more about the environmental impacts.

1 Source: Kaack et al. “Aligning artificial intelligence with climate change mitigation.” Nature Climate Change. Volume 12, June 2022.

2 Source: Kaack et al., June 2022.

3 Source: Saenko, Kate. “Is generative AI bad for the environment? A computer scientist explains the carbon footprint of ChatGPT and its cousins.” The Conversation. 23 May 2023.

Related blogs

+ 4 takeaways from EmTech Digital's AI conference

+ The end of the SaaSacre and the rise of generative AI

+ Tap into the AI revolution with WisdomTree

Important Risks Related to this Article

Important Information

Marketing communications issued in the European Economic Area ("EEA"): This document has been issued and approved by WisdomTree Ireland Limited, which is authorised and regulated by the Central Bank of Ireland.

Marketing communications issued in jurisdictions outside of the EEA: This document has been issued and approved by WisdomTree UK Limited, which is authorised and regulated by the United Kingdom Financial Conduct Authority.

WisdomTree Ireland Limited and WisdomTree UK Limited are each referred to as "WisdomTree" (as applicable). Our Conflicts of Interest Policy and Inventory are available on request.

For professional clients only. The information contained in this document is for your general information only and is neither an offer for sale nor a solicitation of an offer to buy securities or shares. This document should not be used as the basis for any investment decision. Investments may go up or down in value and you may lose some or all of the amount invested. Past performance is not necessarily a guide to future performance. Any decision to invest should be based on the information contained in the appropriate prospectus and affer seeking independent investment, tax and legal advice.

The application of regulations and tax laws can offen lead to a number of different interpretations. Any views or opinions expressed in this communication represent the views of WisdomTree and should not be construed as regulatory, tax or legal advice. WisdomTree makes no warranty or representation as to the accuracy of any of the views or opinions expressed in this communication. Any decision to invest should be based on the information contained in the appropriate prospectus and affer seeking independent investment, tax and legal advice.

This document is not, and under no circumstances is to be construed as, an advertisement or any other step in furtherance of a public offering of shares or securities in the United States or any province or territory thereof. Neither this document nor any copy hereof should be taken, transmitted or distributed (directly or indirectly) into the United States.

Although WisdomTree endeavours to ensure the accuracy of the content in this document, WisdomTree does not warrant or guarantee its accuracy or correctness. Where WisdomTree has expressed its own opinions related to product or market activity, these views may change. Neither WisdomTree, nor any affiliate, nor any of their respective officers, directors, partners, or employees accepts any liability whatsoever for any direct or consequential loss arising from any use of this document or its contents.