JUniper
PUBLISHERS
key to the Researchers

A Short Overview of Rebound Effects in Methods of Artificial Intelligence

Martina Willenbacher[1,2]*, Torsten Hornauer[2] and Volker Wohlgemuth[2]

[1]*Leuphana University of Lüneburg, Institute for Environmental Communication, Germany*

[2]*University of Applied Sciences HTW Berlin, School of Engineering - Technology and Life, Berlin*

*Corresponding author: Martina Willenbacher, Leuphana University of Lüneburg, Institute for Environmental Communication, Universitätsallee 1, 21335 Lüneburg, Germany & University of Applied Sciences HTW Berlin, School of Engineering - Technology and Life, Treskowallee 8, 10318 Berlin

Abstract

Artificial intelligence (AI) is one of the pioneering technologies of the digital revolution in terms of the already existing and potential application areas in companies. On the technical side, this paper deals with the energy requirements for the use of artificial intelligence methods and the identification of efficiency approaches. Productivity increases often lead to increased energy demand, which contradicts sustainability in terms of reducing CO_2 emissions. Therefore, it is investigated to what extent rebound effects can reduce the potential for energy savings in relation to artificial intelligence methods and what are the main factors of the resulting CO_2 emissions.

Keywords: Artificial intelligence; Rebound effect

Introduction

Methods of artificial intelligence have gained enormous popularity in recent years, also due to increasingly cheap storage media and graphical computing units. In particular, artificial multi-layered neural networks have received a lot of media attention in the last ten years and also [1] have achieved a sharp increase in popularity in the scientific field. Especially in the field of artificial neural networks, enormous amounts of data are processed based on nonlinear functions e.g., to enable the system to autonomously find relationships within the data set [2].

The more extensive the training of the AI, the more complex the calculations and the associated hardware requirements, especially regarding the tensor processing unit (TPU), but also the GPU. For the calculation of the algebraic basics, these processors are better suited than a conventional CPU [3]. When using artificial intelligence methods, most of the energy used is not required by the processing of the matrix multiplications themselves, but by accessing the storage areas and the associated rewriting of the data to be processed. Using the example of a CPU-based 8-core server CPU on a 45-nanocenter infrastructure, it was shown that both addition and multiplication processes require only a fraction of the energy of data access. Even the comparatively complex multiplication of floating-point numbers with a size of 32 bits requires little more than a third of the energy of memory access to 8 KB in the processor's cache. If the data is in the main memory, the energy requirement increases by a factor of approx. 1000 [4] (Figure 1).

With increasing size of the training data, the power consumption increases exorbitantly simply by the necessary rewriting of the data.

Energy Consumption and Global Warming Potential of AI

If the increase in efficiency in the energy consumption of one sector leads to the fact that the savings achieved by the lower energy demand are reinvested elsewhere to consume more energy in another way, this is called the rebound effect. The strength of the energy consumption associated with AI processes and the greenhouse gas potential essentially depends on the question of how strongly the purchased electricity is composed of fossil primary energy sources and how efficiently the hardware works for the respective calculation [5]. In addition to transport and extraction, the manufacturing processes themselves have the potential to exacerbate regional conflicts or [6] release toxic chemicals into the environment (Figure 2).

Energy Demand and Global Warming Potential of AI Applications

Strubel et al. [8] from the University of Massachusetts estimated, among other things, the energy requirements, the Co_2 equivalent and the approximate computing costs for cloud providers to train several popular [8] natural language programming models. Particularly noteworthy were the values for a Transformer

model. The basis of this model was the TensorFlow library "Tensor2Tensor" [9,10]. A model with 213 million parameters was used for the Weighting. Neural architecture search (NAS) was used in the existing transformers architecture as an additional feature. For the calculations, a TP integrated Uv2 (tensor processor) chip was used to train the AI. The TPU needed 10 hours [11]. Strubel et al. [8] calculates that when using 8 GPUs (Nvidia Tesla P100), this would correspond to a training time of just over 274 hours. An exponential trend in machine learning in terms of the demand for computing capacity through relatively popular sets has been observed since about 2012.

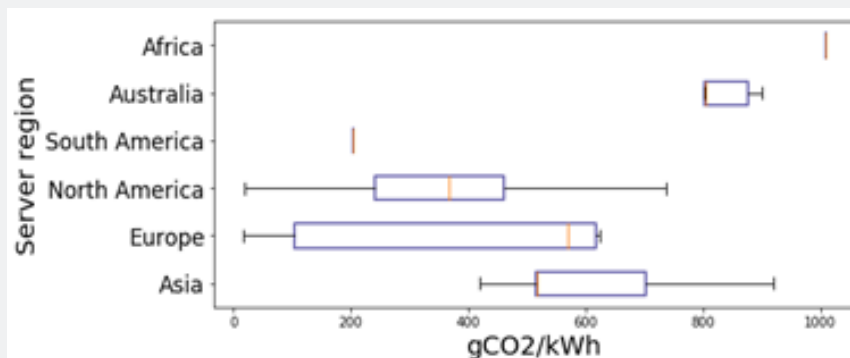**Figure 1:** Energy and time consumption for data access and movement [4].

Figure 2: Quantification of machine learning carbon emissions [7].

Before 2012, it was not common for GPU to be used for computations. The further back in time you look, the more relevant the price differences for RAM and persistent memory may have been. In addition, the development of machine learning may have been accelerated by more powerful algorithms and by the popularity of economic applications.

Green AI and Digitalization

In addition to the enormous environmental impacts described above, the resource requirements of current AI models also represent another critical trend: In orders of magnitude beyond the Evolved Transformer, the costs for electricity and hardware (or cloud service rental) can reach seven or eight figures. The Green AI concept pursues the approach of bringing not only the improvement of a performative best, but also the efficiency with which a certain result was achieved, into the scientific discourse on AI. The weighting could thus be composed of the number of examples used (E), the size of the training data set (D) and the number of AI trainings by hyperparameter experiments (H) and thus scale the result (R) or the accuracy of the AI [12].

$$\mathrm{Cos}\,t(R) \propto E \cdot D \cdot H$$

The resulting weighted costs place a strong focus on efficient AI, as the parameters E, D and H are directly reflected in the computational effort. All in all, Green AI is trying to establish a more complex approach. The focus should not only be on the accuracy of an AI, but also on additional factors such as money, energy consumption and, as a result, the environmental impact.

Discussion and Conclusion

Just as Green AI seeks to convey a more holistic picture of the AI community's goal, from a sustainability perspective, it is necessary to expand efficiency to levels such as consistency and sufficiency. The proportion of recyclable components should be as high as possible. Ideally, the same processed resources are recycled and reused repeatedly for new products [13]. Since intangible goods are created in software development, this logic can only be transferred to the product indirectly through concepts such as frameworks. Rather, software without protection mechanisms is only bound to the hardware requirements in terms of its duplicability. The use of artificial intelligence also raises the question of whether techniques are used for their own sake or whether innovations should be seriously promoted. Rebound effects can also be used as a synonym for growth ambitions. Of course, this is also the case if economic growth is to be stimulated with the help of AI. If a society and its economic system are dependent on growth, there will also be rebound effects. In combination with an efficient circular economy and conscious consumer behavior, the effects of rebound effects can potentially be reduced.

Innovative technologies such as AI can make meaningful contributions – at best if they are used in a targeted manner. More digitization does not necessarily lead to more efficiency.

References

1. Bianchini S, Müller M, Pelletier P (2009) Deep Learning in Science. University of Strasbourg.

2. Russell S, Norvig P (2012) Artificial Intelligence - A Modern Approach. Pearson Deutschland GmbH.

3. Wang Y, Gu-Yeon W, Brooks D (2019) Benchmarking von TPU-, GPU- und CPU-Plattformen für Deep Learning. Johannes A. Paulson School of Engineering and Applied Sciences, Harvard Universität, USA.

4. Horowitz M (2014) The energy problem of computing (and what we can do about it). International Solid-State Circuits Conference, San Francisco, USA.

5. Fritsche UR (2007) Greenhouse gas emissions and reduction costs of nuclear, fossil and renewable electricity supply, Oeko-Institut e.V.

6. Marschall L, Holdinghausen H (2007) Rare Earths - Contested Raw Materials of the High-Tech Age. ekom verlag, Munich, Germany.

7. Lacoste A, et al. (2009) Quantifizierung der Kohlenstoffemissionen des maschinellen Lernens. University of Montreal, Canada.

8. Strubel E, et al. (2019) Energy and Policy Considerations for Deep Learning in NLP. Department of Information und Informatics, University of Massachusetts, Amherst, USA.

9. Github-Repository (2017-2021).

10. Vaswani A, et al. (2017) Attention is all you need. Google Brain.

11. Also DR, Liang C, Le QV (2019) Der Evolved Transformer.

12. Schwartz R, et al (2019) Green AI. Allen Institute of AI.

13. Geissdoerfer M, et al. (2017) Die Kreislaufwirtschaft - Ein neues Nachhaltigkeitsparadigma? Journal of Cleaner Production 143.

Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats
  (Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission
https://juniperpublishers.com/online-submission.php