
Evaluating the carbon footprint of an AI program

2022-2023

1 Carbon footprint of a machine learning program

The objective of this session is to study the environmental impacts of an artificial intelligence program.

The Green Algorithms tool enables to compute the electricity consumption and the carbon footprint of executing a machine learning program.

The electricity consumption is computed as follows:

$$C_{total} = runtime \times \left(\sum_{c \in cores} (P_c \times usage_c) + P_{memory} \right) \times PUE \times PSF$$

with:

- *runtime* the runtime in *hours*
- P_c the power draw of a core c (CPU or GPU) in *Watt*. The actual power draw being unknown, Green Algorithms uses the CPU or GPU TDP, which is considered to be an estimation of the average power draw for the equipment.
- $usage_c$ the use rate for core c , between 0 and 1
- P_{memory} the power draw of the memory
- PUE the power usage effectiveness of the datacenter, over 1. By default, the PUE is the average global PUE: 1,67.
- PSF is the number of times that the training was done, by default 1

Then the carbon footprint is computed as follows:

$$CarbonFootprint = C_{total} \times CI$$

with:

- CI the carbon intensity of the electricity from the region considered

We consider the training parameters in table 1.

In the following questions, you will use Green Algorithms.

1. Calculate the electricity consumption and the carbon footprint associated with the reference scenario.

Runtime	190 h
Number of CPUs	4
Type of CPU	Xeon E5-2683 v4
Number of GPUs	4
Type of GPU	Tesla V100
Memory available per GPU	32GO
Location of the server	Orsay, France

Table 1: Information about model training

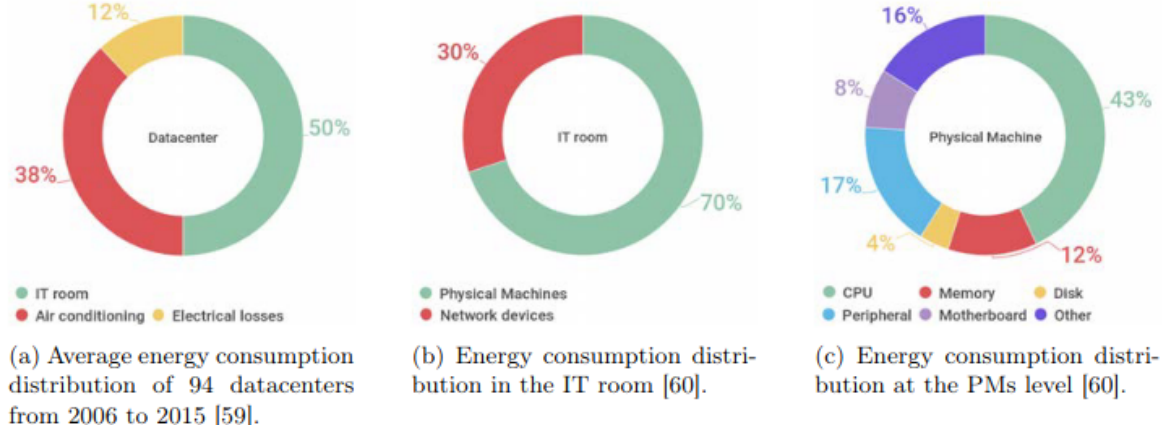


Figure 1: Distribution of electricity consumption in data centers. Physical machines = servers. Source: David Guyon, 2018

2. The default PUE is 1. How does the carbon footprint change if you use a datacenter with a PUE of 1.2? (reset the PUE to its default value afterwards)
3. As a reminder, the PUE is calculated as follows:

$$PUE = \frac{\text{total consumption of datacenter}}{\text{IT equipment consumption}}$$

A significant portion of a data center’s non-IT consumption comes from air conditioning. What does this tell you about the relevance of using an average PUE?

4. If we assume that the electricity consumption of the datacenter follows the distribution of Figure 1, what proportion of the electricity consumption has been studied with Green Algorithms? (be careful however because in this figure, there are only CPUs so it is not really transposable to our study case)
5. What advantages and disadvantages do you see in using such a tool rather than a python package like CodeCarbon?

2 Taking into account the life cycle of equipment

1. If we assume that the distribution of the carbon footprint of the servers in the datacenter follows the distribution of Table 2, what proportion of the carbon footprint was studied with Green Algorithms?

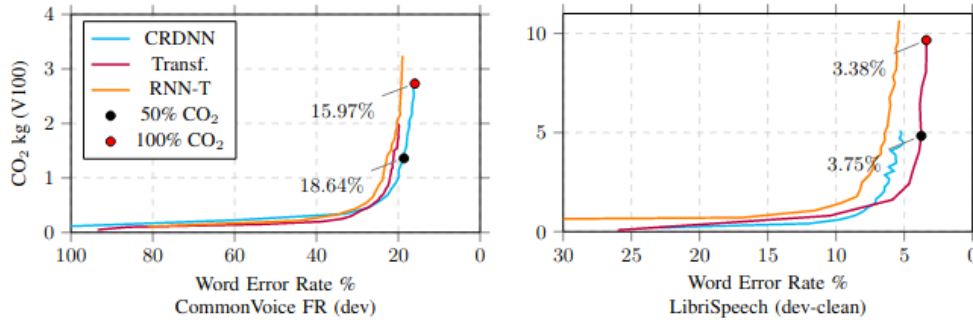


Figure 2: CO₂ emitted in kg (in France) by different E2E ASR models with respect to the word error rate (WER) on the dev sets of LibriSpeech and CommonVoice. The curves exhibit an exponential trend as most of the training time is devoted to slightly reduce the WER. The black and red dots indicates the WER obtained with 50% and 100% of the emitted CO₂. On LibriSpeech, 50% of the carbon emissions have been dedicated to reach SOTA results with an improvement of 0.37%.

Figure 2: Error rate vs carbon footprint. Source : T. Parcollet et al, 2021

2. How do you think the carbon footprint changes if the servers are on a cloud rather than local? Test your hypothesis with an Azure cloud in "France Central" to have a similar location.
3. Does the comparison of cloud vs local servers with Green Algorithms allow to conclude on the environmental interest of the cloud compared to local servers?
4. Compare the resulting footprint with an older model (NVIDIA P100 for example, from 2016) and a newer, more specialized model (TPU v3 for example, from 2018). What do you conclude?
5. What information would you need to estimate the carbon footprint including equipment manufacturing?

Equipment	gCO ₂ e	%
Computation servers (manufacturing)	1.16	29
Other servers (manufacturing)	0.38	10
Computation servers (usage: électricity)	1.79	45
Other servers (usage: électricity)	0.64	16

Table 2: Carbon footprint of servers in the GRICAD infrastructure, for one hour.core of computation. Source: Berthoud et al, 2020

3 Carbon footprint / accuracy trade-off

1. If we had stopped training after 3 days, how much emissions would we have saved?
2. Consider Figure 2 which presents the error rates during training of speech recognition systems for two different corpora (datasets). The authors of the corresponding paper calculated the error rates of the systems when 50% and 100% of the final carbon footprint was emitted. What are the accuracy gains for each curve? When do you think it is reasonable to stop training such a system?

4 Training vs inference

Here we will consider the recently created BLOOM language model. Information about training and inference is available in table 3. We will assume that the PSF of the training is 2, and that the CPU consumption is negligible.

1. Calculate the carbon footprint of running this model.
2. At the time of writing, the model had been in use for about 18 days, with the consumption of using the model emitting about 19kgCO₂e per day. What is the associated carbon footprint for inference?

Runtime	118 days (2 832 hours), 5 hours, 41 min
Number of GPUs	416
Type of GPU	Nvidia A100
Memory available per GPU	80GO (33 280 in total)
Location of the server	Orsay, France
PUE	1.2

Table 3: Information about model training

5 Green or carbon-less electricity?

1. How do you think the carbon footprint evolves if the servers are located in Poland? And in Sweden? Test your assumptions.
2. According to a Google research paper, Google’s renewable energy purchases further reduce the impact to zero ¹. What do you think of this statement?

¹<https://arxiv.org/pdf/2204.05149.pdf>