The Environmental Cost of AI: Incentivizing Efficient Search Through Computational Penalties

Sayeh GHOLIPOUR PICHA, *sayeh.gholipour-picha@grenoble-inp.fr*
Dawood AL CHANTI, *dawood.al-chanti@grenoble-inp.fr*
Alice CAPLIER, *alice.caplier@grenoble-inp.fr*

TL;DR

To promote awareness and integrate the principles of Green AI, we propose presenting CO2e computation as a compulsory component of every AI training process, incorporating it as a key factor in the evaluation of research outcomes, rather than leaving it as a voluntary practice.

Some Statistics

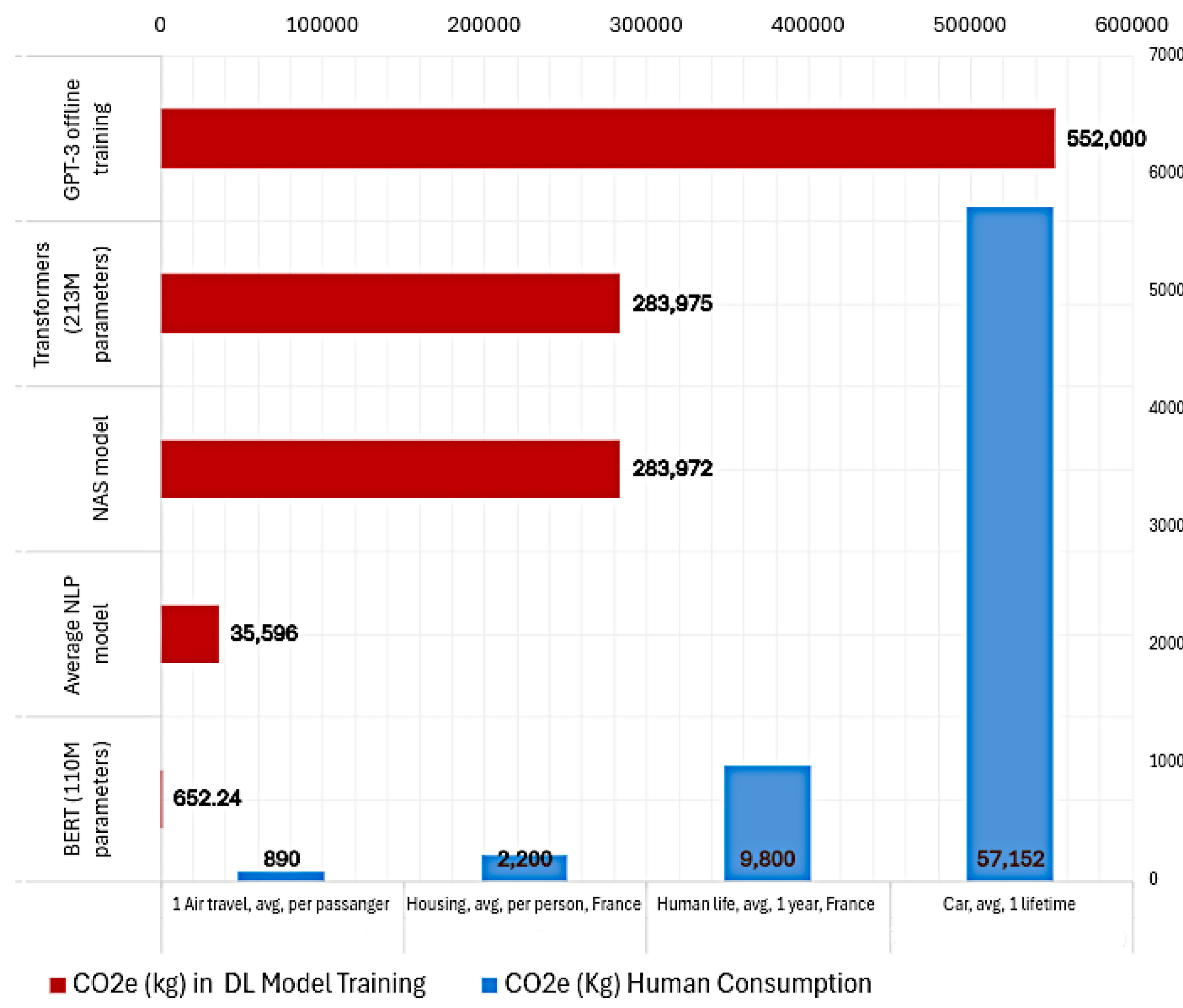
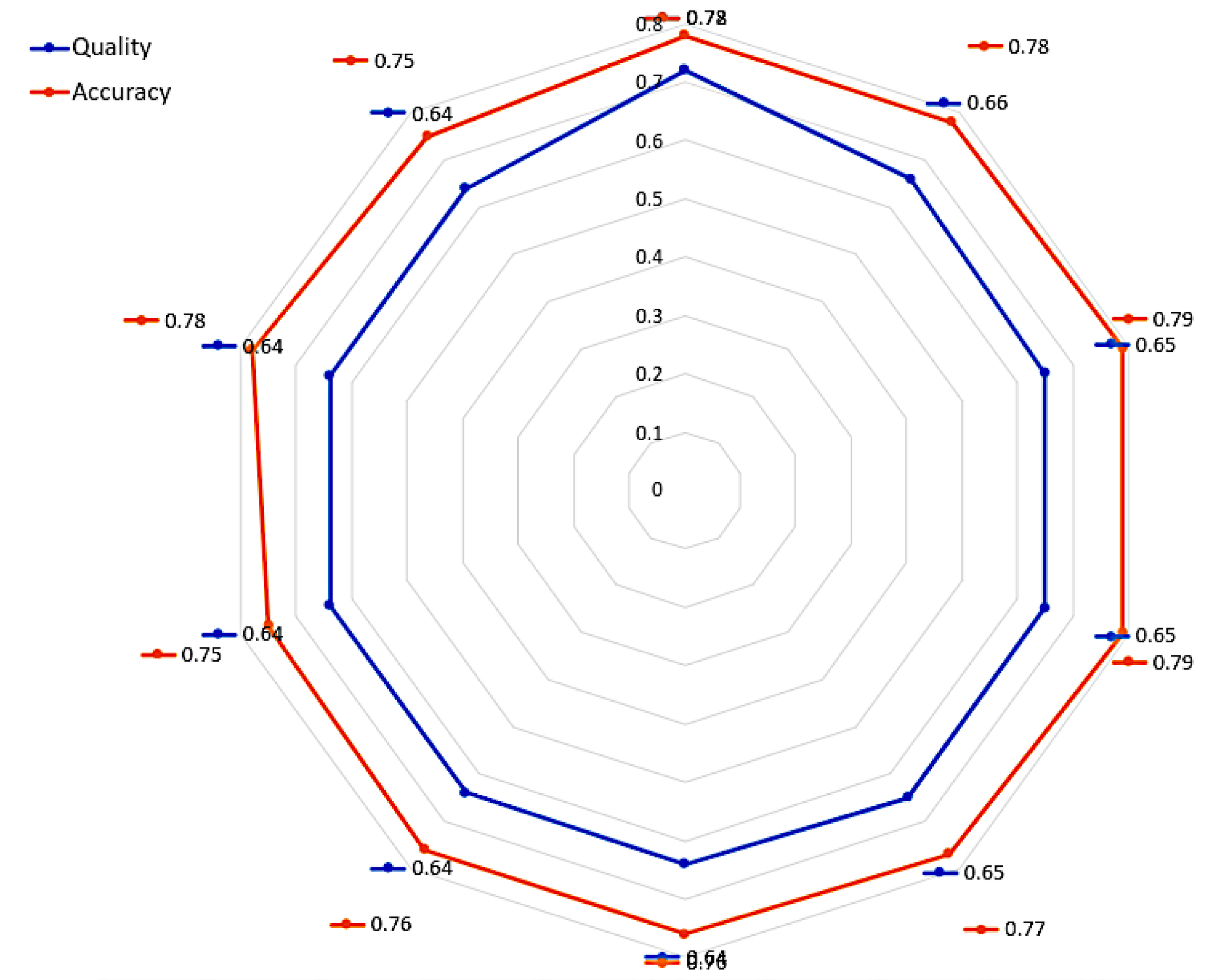


Figure 1: CO2 emissions from human consumption are presented in blue (Ministry of Ecological Transition, 2024), and emissions from AI training are shown in red (AIAAIC, 2024), (Strubell et al., 2019).

- The growth of transformers in deep learning has caused a sharp rise in carbon emissions during training, excluding other environmental costs like data storage and cooling systems.
- A 2021 evaluation showed that training GPT-3 (offline version 2021) produces CO2 emissions 10 times greater than the lifetime emissions of a car. With newer models like GPT-4 using online training, emissions have increased significantly.
- Google AI reported 14.3 billion kilograms of CO2 emissions in 2023, largely due to large-scale models like Gemini.



	Submission1	Submission2	Submission3	Submission4	Submission5	Submission6	Submission7	Submission8	Submission9	Submission10
Accuracy	0.78	0.78	0.79	0.79	0.77	0.76	0.76	0.75	0.78	0.75
Quality	0.72	0.66	0.65	0.65	0.65	0.64	0.64	0.64	0.64	0.64

Figure 2: Results of a visual object tracking challenge: Top 10 submissions out of 80 participants. 9 of these submissions were quite similar, showing no significant improvements over one another.(VOT2024,)

- **Active Competitions:** There are approximately 200 active computer vision competitions on Kaggle.
- **Submission Trends:** On average, each competition receives around 10 submissions, with only minimal performance improvements between them.
- **Stagnant Progress:** The results are no longer advancing significantly by adopting new methods.
- **Trial-and-Error Efforts:** Each submission often involves multiple training cycles, relying heavily on trial and error to fine-tune hyperparameters for slight gains.
- **Lack of Efficiency Evaluation:** Participants are not assessed on the number of training iterations they perform to achieve their final results, overlooking efficiency as a criterion.

Question?

Proposal

1. **Integrating CO2e Metrics**
   - Assessment of CO2 equivalent (CO2e) emissions alongside performance metrics (e.g., accuracy, F1 score). This metric could standardize CO2 evaluation across different research works and ensure it becomes part of the evaluation process during paper reviews and competitions.
   - Encourage conferences and journals to require CO2 estimates in submissions, similar to how runtime or parameter count is often reported.
2. **Adaptive Carbon-Aware Training Strategies**
   - Develop carbon-aware early stopping mechanisms. Instead of stopping based only on performance levels, integrate a threshold based on cumulative CO2 emissions.
   - Implement algorithms that dynamically adjust training hyperparameters to minimize emissions without sacrificing significant performance.

References

AIAAIC (2024). Chatgpt training carbon emission.

Ministry of Ecological Transition (2024). Chiffres clés du climat 2024: Carbon footprint and territorial emissions.

Strubell, E., Ganesh, A., and McCallum, A. (2019). Energy and policy considerations for deep learning in nlp.

VOT2024. https://www.votchallenge.net/vots2024/.

Symposium pour l'électronique & le numérique durables

Le 12 décembre 2024, Grenoble