Deep learning and ecology: what practices, what values for the neuroimaging scientific community?

Nathan Lemmers 1,2, Marie Jacqué 3, Julien Lefevre 2
1,2 INSA Toulouse, 2 Aix-Marseille Université, INT 3 Aix-Marseille Université, LPED.

The investigation we are proposing took place as part of an engineering school internship within the UNITAE ecological redirection project [1] at the Neuromarseille institute. This project is part of a global movement in higher education and research (e.g. [2] ; [3] in computer science and [4] in neuroscience) which aims at 1) quantifying the impact of research activity and 2) through a reflexive approach, identifying the issues at stake within different epistemic communities in order to envisage their future in relation to planetary limits.
In this context, part of this study focused on the use of deep/machine learning methods (AI in the following) in neuroscience laboratories, and more generally in medical research. In this we proposed exchanges with those involved in the research of this field. The goal is to analyze scientists' practices from two angles: the first aims to quantify precisely the energy consumption of deep learning algorithms [5]; the second questions the motivations that might drive researchers to rethink the production of knowledge coming from AI.

1) Quantifying numerical computations

In order to think about and work on reducing the laboratory's energy consumption, the first step must be precise quantification.
With this in mind, the first stage of the project was to work on a way of measuring the power consumption of deep/machine learning methods. Since neuroscience researchers mainly work on shared servers, and not all of these have built-in power meters, it is necessary to work with digital power meters. There are already many available [6], so we tested the main ones, to evaluate the methods and understand why some were not relevant in our situation.
This work led to the design of our own module, adapted to the needs of the neuroscience laboratory, and more specifically to the researchers working there. This tool, although inspired by the relevant points of the other methods, nonetheless brings a number of new features to enhance measurement precision, with real-time carbon intensity measurement for example. In addition, the problems of other tools in the context of a shared server have been highlighted, so we can look forward to a future update.
In conclusion, our tool can enable the laboratory's researchers to accurately measure the impact of their numerical computations. In a next step we are planning to evaluate the tool on a larger scale and over a longer period before a more systematic deployment, possibly with the GES1point5 tool [7]. Measuring the consumption of these new methods, and their impact on climate change, will provide us with a solid basis for future democratic choices.

2) Understanding the challenges of AI for the research activity

Quantifying the impact of AI methods is not enough to determine exactly how much reduction is needed. Indeed, it is important to reveal and understand the dependencies and attachments of our community to connectionist methods, to realize the consequences of a possible reduction in consumption or usage.

To this end, we have exchanged views with actors of the neuroscience research community to assess their frequency of use of AI. To this end, we conducted a qualitative pre-survey of 8 semi-directive interviews with members of three neuroscience laboratories. As the AI sector is particularly male-dominated, we only interviewed men at the moment. These discussions lasted between 30 minutes and 2 hours, and enabled us to collect the different contexts in which AI methods are used, to understand the associated changes of practice, and to analyze the different types of discourse. On the basis of a content analysis, we constructed three typical ideals [8] of the link between AI and ecology. The ecological discourse, although rare, advocates a global reduction in our emissions, even if this impacts on the efficiency and progress of research. The second, on the contrary, is techno-solutionist, convinced that these technological advances are necessary, useful and effective, and will enable us to solve all future problems. The last type of discourse is aware of the inequalities that artificial intelligence reinforces or brings to light, via biased data that favors certain types of population. This last type is concerned with the thoughtful use of AI.
Drawing on the article "The values encoded in machine learning research" by A. Birhane [9], we noted the occurrences of the values that our actors linked to AI, and were able to notice numerous concordances, via keywords such as performance, efficiency, usefulness and so on. In addition, we noted a real gap in the mention of inequality generated by AI, which we associated with our last type of discourse.

It is important, however, to recall the possible biases of our study. The corpus of discourse is constructed from interviews conducted solely with men, who expected to talk about ecology, and who knew us before the interviews began. This partly explains the over-representation of ecological and techno-solutionist discourse. In both cases, the main finding is that the values of efficiency, productivity and utility seem to be shared by the scientific community. These initial results show that some researchers are integrating environmental concerns into their use of technology. However, questions concerning the place of AI in the production of scientific knowledge, and the scientific and ethical value of the data produced by AI, remain to be explored. Nor can an analysis of the links between AI and ecology avoid a reflection on the responsibilities of the scientific community with respect to the increased production of data on the one hand, and its dissemination on the other, in a paradoxical context of open science.

Conclusion

To sum up, our study within what we can consider an epistemic community, has enabled us to make progress on very recent subjects through an analysis of the current evolution of this community. Our work has provided the laboratory with the tools to measure its energy consumption and consider reduction strategies. We have also initiated a reflection on the evolution of research values and practices, and how our current system could integrate or rethink scientific practice in the Anthropocene era.

Bibliographie

[1] https://zenodo.org/records/10122489

[2] Ben-Ari, T. How research can steer academia towards a low-carbon future. Nat Rev Phys5, 551–552 (2023). https://doi.org/10.1038/s42254-023-00633-9

[3] Florence Maraninchi. 2022. Let us not put all our eggs in one basket. Commun. ACM 65, 9 (September 2022), 35–37. https://doi.org/10.1145/3528088

[4] Anne E Urai Clare Kelly. Point of View: Rethinking academia in a time of climate crisis eLife 12:e84991. (2023) https://doi.org/10.7554/eLife.84991

[5] Bouza Heguerte, Lucia, Aurélie Bugeau, and Loïc Lannelongue. "How to estimate carbon footprint when training deep learning models? A guide and review." Environmental Research Communications (2023).

[6] M. Jay, V. Ostapenco, L. Lefèvre, D. Trystram, A.-C. Orgerie, and B. Fichel, "An experimental comparison of software-based power meters: focus on CPU and GPU," in IEEE/ACM international symposium on cluster, cloud and internet computing, May 2023.

[7] https://apps.labos1point5.org/

[8] Max Weber, L'objectivité de la connaissance dans les sciences et la politique sociales (1904)

[9] A. Birhane, P. Kalluri, D. Card, W. Agnew, R. Dotan, M. Bao. 2022. The Values Encoded in Machine Learning Research. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22). Association for Computing Machinery, New York, NY, USA, 173–184. https://doi.org/10.1145/3531146.3533083

Biographie

Nathan Lemmers est étudiant en 5ème année de l'INSA Toulouse dans la spécialité informatique et réseaux. En 2021, il a réalisé un échange avec l'Université métropolitaine d'Oslo. En 2023 il a réalisé son stage de 4ème année à l'Institut de Neurosciences de la Timone, dans le cadre du projet Unitae « Quelles neurosciences à l'anthropocène ? ».

Marie Jacqué est maîtresse de conférences HDR en sociologie à Aix-Marseille Université au Laboratoire Population Environnement Développement. Elle mène ses recherches sur les pratiques, usages et formes de gestion des ressources naturelles en Méditerranée. Ses travaux portent également sur la diffusion des savoirs scientifiques par l'intermédiaire des pratiques d'éducation à l'environnement.

Julien Lefevre est maitre de conférences HDR en informatique à Aix-Marseille Université à l'Institut de Neurosciences de la Timone. Ses recherches portent sur la modélisation géométrique du cerveau avec des travaux théoriques sur l'analyse spectrale des surfaces et des applications en neurosciences. Depuis trois ans il réoriente une partie de ses activités sur l'impact environnemental du numérique.