Towards a Methodology and Framework for AI Sustainability Metric

Some authors

Recently, we are witnessing truly groundbreaking achievements using AI models, such as the much talked about generative large language models, and a wide range of application, with a tremendous potential to accelerate scientific discovery and to enhance productivity. At the same time, AI models and their use are growing at an exponential pace, and consequently, their environmental footprint implications are staggering. There is an urgent need for the community, including data scientists, businesses, academia, government, and practitioners, to come together and conduct a meaningful conversations about the environmental cost of AI. For this, the authors of this article posit that we ought to develop an agreed upon methodology, and framework to quantify AI efficiency across the life cycle. In particular, we factor in recent advances bringing to the front the concept of foundation models, promoting model re-use, as a strategy to drive efficiencies. We claim that a methodology and framework for Sustainable AI must provide the insight necessary for the different stake holders, including data scientists, and practitioners, to make informed decisions, across the model life cycle, factoring in trade-offs such as accuracy, time, environmental impact, and societal benefits. Such a methodology and framework must be consistent with existing standards used in other domains such as the GHG protocols and accompanying guidance documents, and LCA standards. Recent work on quantifying the environmental cost of AI, falls short in multiple areas, such as, that it either limits the discussion to the examples perused by the businesses behind the technology under examination, and are not easily generalizable; they are ignoring the cost of maintaining an AI model *after* the initial training and deployment phase, including model distillation, fine-tuning, and re-training, and they have no way to factor in the benefits of model re-use, which is at the core of the promise of foundation models, thus injuring our ability to quantify the benefits of life cycle strategies. In this paper, we analyse existing limitations of proposed approaches for quantifying the environmental cost of AI. We take a first step in proposing a methodology that addresses these limitations. We exemplify the concepts, using a complex production scale example of LLM across the life cycle. We analyse the requirements for a framework that will produce the necessary telemetry and computational data to support the various stake holders in the decision making process, throughout the life cycle, and we list opportunities for further research in this critical area.

Categories and Subject Descriptors: I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism—*Animation*; I.3.5 [**Computer Graphics**]: Computational Geometry and Object Modeling—*Physically based modeling*

General Terms:

Additional Key Words and Phrases:

1. BACKGROUND AND MOTIVATION

Artificial Intelligence (AI) is one of the fastest growing technology domains, involving academic research, businesses, and users. The enormous investment in AI led to groundbreaking applications in a diverse set of areas. AI is used for accelerating the discovery of drugs ([**?**]), driving efficiencies at work (e.g., [Puri et al. 2021]), discovering new materials towards renewable storage (e.g., [Zitnick et al. 2020]), and more. [Rous 2008]

APPENDIX

REFERENCES

Ruchir Puri, David S. Kung, Geert Janssen, Wei Zhang, Giacomo Domeniconi, Vladimir Zolotov, Julian Dolby, Jie Chen, Mihir Choudhury, Lindsey Decker, Veronika Thost, Luca Buratti, Saurabh Pujar, Shyam Ramji, Ulrich Finkler, Susan Malaika, and Frederick Reiss. 2021. CodeNet: A Large-Scale AI for Code Dataset for Learning a Diversity of Coding Tasks. (2021).

Bernard Rous. 2008. The Enabling of Digital Libraries. *Digital Libraries* 12, 3, Article 5 (July 2008). To appear.

C. Lawrence Zitnick, Lowik Chanussot, Abhishek Das, Siddharth Goyal, Javier Heras-Domingo, Caleb Ho, Weihua Hu, Thibaut Lavril, Aini Palizhati, Morgane Riviere, Muhammed Shuaibi, Anuroop Sriram, Kevin Tran, Brandon Wood, Junwoong Yoon, Devi Parikh, and Zachary Ulissi. 2020. An Introduction to Electrocatalyst Design using Machine Learning for Renewable Energy Storage. (2020).

Received ; accepted