

RESTART – Energy Efficient training at the Edge

Background

The success of AI is partly due to the usage of the energy-hungry computational infrastructures constantly deployed to train large AI models and provide AI-based services, consuming huge quantities of data, typically stored in the Cloud. Depending on the type of source of energy adopted, the carbon footprint related to AI can represent a significant entry in the overall CO₂-emissions budget. Beyond improving on the type of source of energy used in the process, it is paramount to rethink the whole AI paradigm by including sustainability by design. The new paradigm called "EdgeAI" can also help in this direction. In EdgeAI, part of the AI process to the Edge of the network, leveraging as much as possible the computational capabilities of Edge devices to extract knowledge from locally collected or generated data. However, the devices at the Edge of the network have limited computational and energy resources. While Edge AI offloads energy consumption from data centers, it impacts on the resource consumption of edge devices. Therefore, to realise such a shift it is necessary to rethink how the training is performed on edge devices, designing new algorithms that embed energy and resource efficiency by design.

Topic description

The topic addresses the challenges connected to training AI models on resource constrained Edge devices. In this line of work, the idea is to explore new methods and to design algorithms to make the training of AI models more resource efficient, considering energy efficiency as the prime target. This might include the design of AI models that are energy efficient by design or to design algorithms and methods to reduce the requirements of AI models in terms of computation (e.g. FLOPs), memory and, in general, the energy required to run them. The expected output can lean more towards methodological contributions (e.g., general purpose efficient training algorithms) or be more focused on systems research (e.g., design, development and test of energy-efficient solutions in real edge devices).

Type of prospect positions

We plan to open positions at the level of Researcher on this topic.

Funding and partnerships

Supported by: **RESTART**: Extended Partnership on Telecommunications of the Future (funded by the National Recovery and Resilience Plan (NRRP), European Union - NextGenerationEU)

Candidate profile

Ideal candidates should have obtained a PhD degree (or have a 3-year professional experience in research) in Computer Science, Wireless Networking, Computer Engineering, Telecommunications Engineering, Mathematics, Physics or closely related disciplines, and a proven track record of excellent scientific publications. Preferably, the PhD should be in Artificial Intelligence, Pervasive Systems, Networking.

Contacts



Chiara Boldrini

chiara.boldrini@iit.cnr.it

Scholar profile

<https://scholar.google.com/citations?user=rCrzxlsAAAAJ>



Lorenzo Valerio

lorenzo.valerio@iit.cnr.it

Scholar profile

<https://scholar.google.com/citations?user=CjoX4voAAAAJ>



Andrea Passarella

andrea.passarella@iit.cnr.it

Scholar profile

<https://scholar.google.com/citations?user=sesKnygAAAAJ>



Marco Conti

marco.conti@iit.cnr.it

Scholar profile

<https://scholar.google.com/citations?user=KniFTD0AAAAJ>