Institut de Recherche en Informatique de Toulouse

Millian POQUET
Maître de conférences
Université Paul Sabatier, IRIT
118 Route de Narbonne, 31062 Toulouse Cedex 9, France
millian.poquet@irit.fr

Exploring the tradeoffs between energy and performance of federated learning algorithms

Context

There is an increasing interest in a new distributed ML paradigm called Federated Learning (FL)[La17], in which nodes compute their local gradients and communicate them to a central server. This centralized server then orchestrates rounds of training over large data volumes created and stored locally at a large number of nodes. This training procedure repeats until some criterion are met. This enables the participating nodes (e.g., IoT devices, mobile phones, etc) to protect their data and solve the data security and privacy issues imposed by law.

This growing demand for FL technology opens new challenges in addition to those that appear in traditional ML. These include (i) communications and energy costs between the nodes and the server; (ii) increased duration of the training period due to correlation between databases.

The ANR project DELIGHT will start in early 2023 to address these challenges. In the context of this internship we will focus on the first challenge. This project will use the Flower federated learning framework for demonstration purpose.

Objective of the internship

The main objective of this internship is to explore the tradeoffs between energy and performance of federated learning algorithms. For this purpose we plan to leverage many parameters of the framework, such as the precision of the floating-point computations or how data is transferred between distributed actors. We will focus on the use case of automatic speech recognition.

Phases of the internship.

- Set up an experimental environment on Grid'5000 to gather performance and energy metrics.
  - Create a use-case related to speech recognition for the Flower framework.
  - Build a reproducible and automated framework for obtaining metrics for this use-case.
- Propose, model and implement the different leverages.
- Explore the impact of the leverages on both energy and performance.

After this internship, there is a possibility to continue working on this subject during a PhD. In the context of the DELIGHT project, a PhD funding is available to work on reproducible FL experiments and simulations. This PhD is scheduled to start on Sept. 2023 just after the internship. This PhD will particularly focus on sustainable FL.

Expected skills and profile

- Performance evaluation skills
- Programming skills (Python)
- A taste for experimental methods (a taste for chocolate is a plus)
- Fluent English or French

Practical details

The internship will take place at IRIT, the largest computer science research institute in Toulouse, France. Our team SEPIA works on resource management on various distributed systems (cloud data-centers, HPC centers, edge architectures, IoT. . .) and is especially interested in ecological transition, notably by reducing energy consumption and CO_2 emissions, by using renewable energy. . .

The internship will be supervised by Millian Poquet and Georges Da Costa in a convivial atmosphere :). A computer and an office will be provided, as well as a monthly internship stipend of 591 €. Internship duration is 5-6 months. You can send us your application (cover letter + resume / short curriculum vitæ) by email to `millian.poquet@irit.fr` and `georges.da-costa@irit.fr`.

Bibliography

- [La17] P. Kairouz, M. Bennis, et al. *Advances and Open Problems in Federated Learning.*