Contribution ID: **33**
Type: **not specified**

Services for Machine Learning applications (part 3 of 3)

Thursday 7 November 2024 14:30 (1 hour)

This session will focus on available ML techniques for distributed training of models, hyperparameter optimization and model service. In particular, starting from a well known use case it will demonstrate:

- How to go from a script, to a docker image training on a single node, to a distributed training setup with multiple nodes

- How to do hyperparameter optimization, which kind of optimizers are available, how to monitor the workloads and how to publish the models

- How to serve models in production, at scale, with a simple http entrypoint or embedding the model in an application

Presenter: CHIORESCU, Raulian-Ionut