

Services for Machine Learning applications (part 1 of 3)

Thursday 7 November 2024 10:00 (1 hour)

This session will introduce the different phases in a ML lifecycle, and how IT services can help in each of the parts. In particular, it will:

- Overview of ML and use cases, containerization and how it helps out in defining single units of computation, isolate custom software environments, and ensure sustainability for reproducible results
- Demo how cloud native environments (Kubernetes and its ecosystem) can help manage those units of computation and scale them out to large amounts of resources
- Provide an example of how to scale out using both on-premises and public cloud resources, and when this might be useful and cost effective

Presenter: ROCHA, Ricardo (CERN)