



Contribution ID: 243

Type: Talk

Zero-overhead training of machine learning models with ROOT data

Wednesday 23 October 2024 14:06 (18 minutes)

The ROOT software framework is widely used in HENP for storage, processing, analysis and visualization of large datasets. With the large increase in usage of ML for experiment workflows, especially lately in the last steps of the analysis pipeline, the matter of exposing ROOT data ergonomically to ML models becomes ever more pressing. This contribution presents the advancements in an experimental component of ROOT that exposes datasets in batches ready for the training phase. This feature avoids the need for intermediate data conversion and can further streamline existing workflows, facilitating direct access of external ML tools to the ROOT input data in particular for the case when it does not fit in memory. The goal is to keep the footprint of using this feature minimal, in fact it represents just an extra line of code in user application. The contribution demonstrates such usage in various examples using different ML training models, also evaluating the performance with key metrics.

Primary authors: Dr PADULANO, Vincenzo Eduardo (CERN); PRANCKIETIS, Kristupas (Vilnius University (LT)); MONETA, Lorenzo (CERN)

Presenter: Dr PADULANO, Vincenzo Eduardo (CERN)

Session Classification: Parallel (Track 5)

Track Classification: Track 5 - Simulation and analysis tools