

Checkpointing for long running Machine Learning Tasks

Thursday, 4 July 2024 15:00 (30 minutes)

The increasing application of Machine Learning (ML) in High Energy Physics (HEP) analysis and reconstruction necessitates the use of GPUs. The extensive runtimes associated with training neural networks make them vulnerable to runtime constraints and failures.

Checkpointing, which involves storing the current state of the training persistently, offers a solution to these challenges. It allows for continuing training at a later time or different location, providing resilience against failures and adherence to time constraints.

Moreover, checkpointing contributes to sustainable computing efforts. For instance, training can be scheduled during periods of abundant renewable energy supply and paused when the supply is limited.

This presentation introduces a Python interface that consolidates common HEP community tools for storing checkpoints and rescheduling tasks. Examples on how to use this tool in combination with different setups like luigi workflow management or htcondor will be shown.

Primary author: EPPELT, Jonas (Karlsruher Institute of Technology (KIT))

Co-author: SCHNEPF, Matthias

Presenter: EPPELT, Jonas (Karlsruher Institute of Technology (KIT))

Session Classification: Plenary Session Thursday