Contribution ID: **191** Type: **Talk without Poster**

Reinforcement learning for automatic data quality monitoring in HEP experiments

Wednesday, 1 May 2024 16:12 (20 minutes)

The application of modern Machine Learning (ML) techniques for anomaly detection in collider physics is a very active and prolific field, with use cases that include the exploration of physics beyond the Standard Model and the detection of faults in the experimental setup. Our primary focus is on data-quality monitoring. Within large experimental collaborations, this anomaly detection task usually relies on a large pool of rotating non-expert shifters. Their goal is to identify detector-related issues that would render data unusable for future physics analysis. The partial automation of these tasks presents an opportunity to ameliorate data collection efficiency and reduce the need for associated person power.

Challenges intensify in scenarios of rapidly changing experimental conditions, such as during the commissioning of a new detector. In this case, for an automated anomaly detection system to be useful it would have to be continuously retrained in an efficient manner. Additionally, optimization for factors beyond data-collection efficiency, such as minimizing unnecessary human interventions, introduces complexities in defining an adequate loss function. To address these challenges, we propose the application of Reinforcement Learning (RL) techniques with human feedback to the task of data-quality monitoring.

This contribution describes a simplified simulated setup designed to study the automation of data-quality monitoring in two regimes: "online" and "offline". The "online" regime addresses real-time detection of issues in the detector during data collection, emphasizing a rapid intervention to enhance future data-collection efficiency. On the other hand, the "offline" regime centers on the classification of previously collected data as either usable or unusable.

This work aims to exploit RL algorithms within these regimes, demonstrating progress in simulations using both multi-agent and single-agent RL techniques. We present the performance obtained with different policies and outline future research directions.

Primary authors: GARCÍA PARDIÑAS, Julián (CERN); Mr DEL PIANTA PÉREZ, Lorenzo (CERN); Mr JANISCH, Maximilian (University of Zürich); Prof. SERRA, Nicola (University of Zürich); Ms JULLIAN PARRA, Olivia (CERN); Dr KLAVER, Suzanne (NIKHEF); Dr LEHÉRICY, Thomas (University of Zürich)

Presenter: Ms JULLIAN PARRA, Olivia (CERN)

Session Classification: 4.1 Pattern recognition, Image analysis & Uncertainty quantification