

fast machine learning for science

Contribution ID: 93 Type: Standard Talk

Accelerated Graph Neural Network Inference on FPGAs for Real-Time Muon Triggering at the HL-LHC

Wednesday 3 September 2025 14:20 (20 minutes)

The High Luminosity upgrade of the Large Hadron Collider (HL-LHC) presents a demanding environment for real-time data processing, with substantially increased event rates requiring faster and more efficient trigger systems. This study explores the deployment of graph neural networks (GNNs) on field-programmable gate arrays (FPGAs) for fast and accurate inference within future muon trigger pipelines. By leveraging the sparse and relational structure of detector data, GNNs enable robust pattern recognition while preserving spatial and topological correlations. We investigate hardware-friendly implementations of GNN architectures, focusing on model compression, parallelism, and low-latency execution; contributing to the broader goal of AI-driven event selection in high-energy physics experiments.

Authors: FIACCO, Davide (Sapienza Universita e INFN, Roma I (IT)); GUSTAVINO, Giuliano (Sapienza Universita e INFN, Roma I (IT)); RUSSO, Graziella (Sapienza Universita e INFN, Roma I (IT)); ERRICO, Martino (Sapienza Universita e INFN, Roma I (IT)); GIAGU, Stefano (Sapienza Università di Roma and INFN Roma (IT)); Dr IPPOLITO, Valerio (INFN Sezione di Roma (IT))

Presenter: FIACCO, Davide (Sapienza Universita e INFN, Roma I (IT))

Session Classification: Contributed talks