Contribution ID: **11** Type: **not specified**

Tree Tensor Network implementation on FPGA

Tuesday 5 November 2024 10:05 (20 minutes)

Starting from the statements of the previous talk, we present the implementation on FPGA of Tree Tensor Networks as binary classifiers, highlighting the possibility of their deployment in high-frequency real-time environments, such as the online trigger systems of HEP experiments. The linear algebra operations needed by TTNs make them easily deployable on FPGAs, which are extremely suitable for concurrent tasks like matrix multiplications and tensor contractions. In this talk, we provide an intuitive description of the inference firmware achieving sub-microsecond latency, together with a projection of the necessary hardware resources for different combinations of TTN hyperparameters and degrees of parallelization. We probe the feasibility of the hardware implementation of TN classifiers for HEP applications, exploring future prospects and further developments, possibly exploiting the usage of AI Engines on AMD Versal boards.

Presenter: BORELLA, Lorenzo (Universita e INFN, Padova (IT))

Session Classification: Prospects on Tensor Networks and Machine Learning