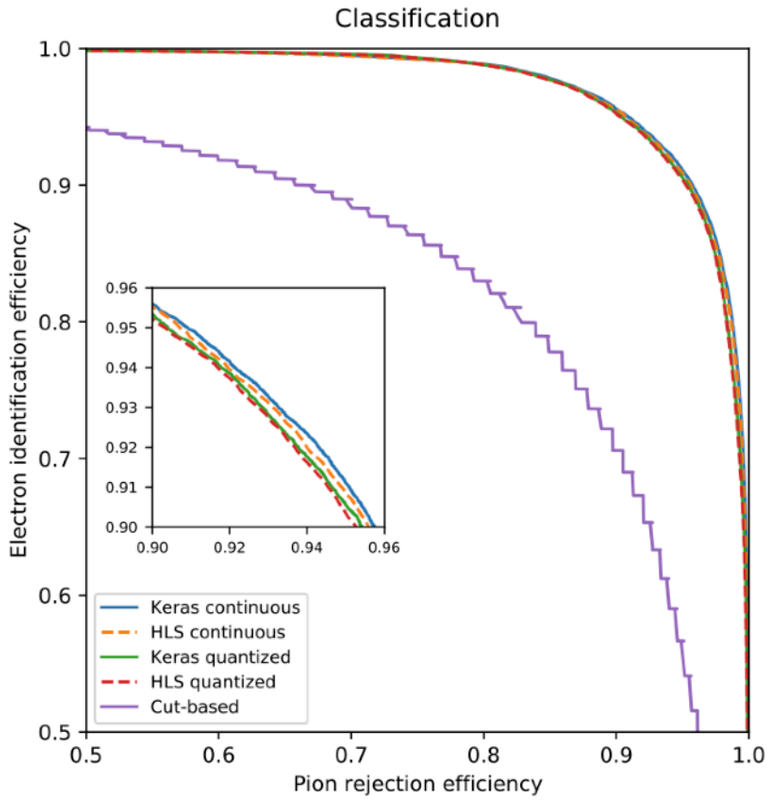
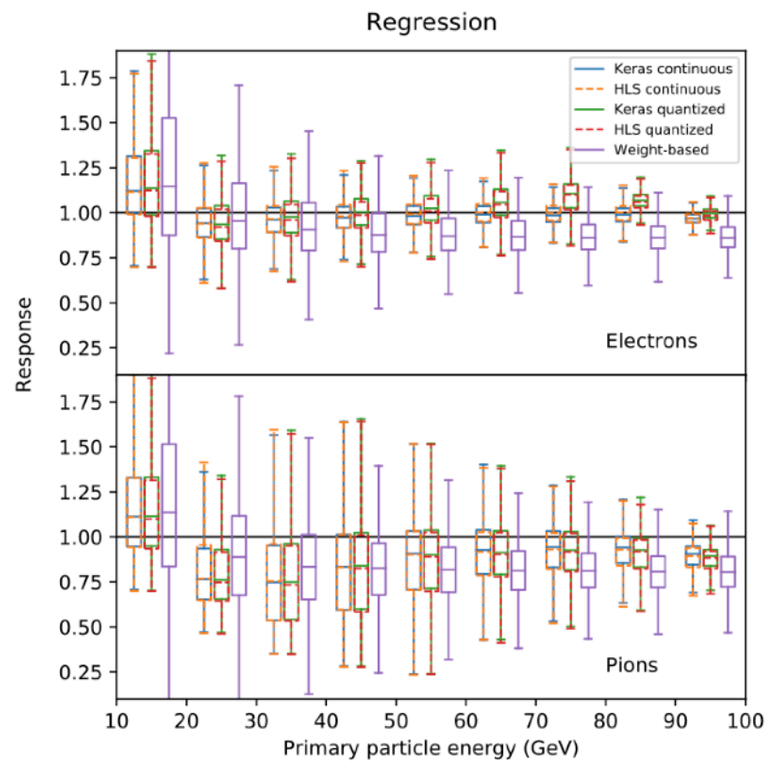
Graph neural networks for FPGAs

- Key challenges for implementing GNNs in FPGAs
  - Number of operations
  - Memory capacity and access
- GarNet[1] addresses these issues
  - No vertex-to-vertex edges
    → Low RAM usage, regular access
  - Nonlinearity in edge weights
    → Encoder / decoder can be shallow
- Algorithm further simplified to fit in FPGA
  - Ideas applicable to other GNNs
  - GarNet HLS implementation available in HLS4ML[2]
- Case study with a toy 3D calorimeter
  - Achieved sub-μs latency for L1T-like physics task





[1] https://arxiv.org/abs/1902.07987
[2] https://fastmachinelearning.org/hls4ml/