Accelerating Hyperparameter Optimization of Deep Learning Models Using Performance Prediction and Quantum Computing

Challange

In CoE RAISE, we have seen that ML can unlock new potential in fields such as high energy physics (HEP), remote sensing, seismic imaging, additive manufacturing, and acoustics. Training Deep Learning (DL) models, however, is no trivial task, especially if the models are large and have many tunable hyperparameters. To tackle this challenge, Hyperparameter Optimization (HPO) can be used to systematically explore the search space of possible hyperparameter configurations. This tends, however, to be a very computationally expensive process as training the target model with a lot of different configurations is usually required.

Solution

Making use of novel approaches for HPO based on performance prediction strategies have potential to drastically speed up the process of improving DL models. We studied the suitability of using such new strategies for MLPF, a neural network for particle flow reconstruction that is being developed in CoE RAISE WP4 in Task 4.1 "Event reconstruction and classification at the CERN HL-LHC".
To do this, we made use of High Performance Computing (HPC) resources to generate a dataset of MLPF learning curves, where we could test the performance predictors. In addition to using classical Support Vector Regression (SVR), a model that has been successfully used for performance prediction tasks in other studies, we collaborated with our colleagues in WP2 to train the performance predictors using an annealing quantum computer, the D-Wave Advantage™ System JUPSI at the Jülich Supercomputing Center. We successfully designed strategies to overcome some of the current limitations of quantum technologies for the task of learning curve prediction in HPO acceleration.

Impact

We demonstrated the strong potential of using performance prediction techniques for HPO, leaving the door open for the use of this technique in later HPO cycles of MLPF. We also showed that, despite the current limitations of quantum computers, it is possible to train SVR models on a quantum annealer while achieving performances comparable to those obtained with classical SVR techniques. This encourages further studies in utilizing hybrid quantum/HPC workflows for HPO as well as in other use-cases.