Contribution ID: **416** Type: **Parallel**

Hyperparameter Optimisation in Deep Learning from Ensemble Methods: Applications to Proton Structure

Deep learning models are defined in terms of a large number of hyperparameters, such as network architectures and optimiser settings. These hyperparameters must be determined separately from the model parameters such as network weights, and are often fixed by ad-hoc methods or by manual inspection of the results. An algorithmic, objective determination of hyperparameters demands the introduction of dedicated target metrics, different from those adopted for the model training. Here we present a new approach to the automated determination of hyperparameters in deep learning models based on statistical estimators constructed from a ensemble of models sampling the underlying probability distribution in model space. This strategy requires the simultaneous parallel training of up to several hundreds of models and can be effectively implemented by deploying hardware accelerators such as GPUs. As a proof-of-concept, we apply this method to the determination of the partonic substructure of the proton within the NNPDF framework and demonstrate the robustness of the resultant model uncertainty estimates. The new GPU-optimised NNPDF code results in a speed-up of up to two orders of magnitude, a stabilisation of the memory requirements, and a reduction in energy consumption of up to 90\% as compared to sequential CPU-based model training. While focusing on proton structure, our method is fully general and is applicable to any deep learning problem relying on hyperparameter optimisation for an ensemble of models.

Secondary track

Author: STEGEMAN, Roy (The University of Edinburgh)

Session Classification: T16

Track Classification: T16 - AI for HEP (special topic 2025)