Contribution ID: **36** Type: **Oral Presentation**

Application of transfer learning to event classification in collider physics

Monday, 21 March 2022 16:30 (30 minutes)

Transfer Learning technique has been successfully applied to many scientific fields such as computer vision, natural language processing, and so on. This presentation reports an enhancement of data analysis in collider physics experiments based on this Transfer Learning technique.

Experimental particle physics aims to understand the fundamental laws of nature using a huge amount of data. In collider physics experiments, each event of data is produced from particle collisions using a high energy accelerator, such as the Large Hadron Collider. The classification of events is quite important in data analysis, where interesting signal events are separated from background events as much as possible.

Deep Learning (DL) technique has been widely used to enhance the performance of event classification by utilizing the huge parameter space of model. However, DL with such huge parameter space requires a large amount of data to maximize performance. In the field of collider physics, training data are typically generated by Monte Carlo simulations based on theories for signal and background processes. However, the simulations cost high computational power to generate a large number of events. Therefore, applying DL technique with a limited amount of data is a key concept for the collider physics experiment.

DL model consists of a stack of layers with non-linear functions. It is considered that their initial part of layers learns local features of data, and then the subsequent layers learn global features. This indicates that knowledge gained while solving one problem, such as extracting local features, can be transferred to different problems which involve common tasks. In the collider physics experiments, there are many data analysis channels for targeting different signal events. If DL can learn common knowledge or features for different data analysis channels, Transfer Learning (TL) technique should work effectively. This technique allows us to avoid training DL models from scratch by re-using weight parameters. If these weight parameters can be re-used for many data analysis channels, we can save a lot of computational power for generating the simulation data.

In this presentation, we report that the event classification can be performed with a high accuracy for different signal events by applying the TL technique. The event classification is typically based on the information of final state particles (objects) from collisions. A re-training with a small amount of data (fine-tuning) is performed to absorb differences of object topologies. For example, the number of objects in the final state is different depending on the signal process. Thus, the DL model needs to work with a variable number of objects and be insensitive to the ordering of objects to apply TL effectively. We propose a DL model to overcome these problems. The proposed model is compared to a simple multilayer perceptron model, which has a similar number of trainable parameters. Technical details of the DL model and limitations of this study are also discussed.

Primary authors: TANAKA, Junichi (University of Tokyo); SAITO, Masahiko (ICEPP, The University of Tokyo); MORINAGA, Masahiro; KISHIMOTO, Tomoe (High Energy Accelerator Research Organization)

Presenter: KISHIMOTO, Tomoe (High Energy Accelerator Research Organization)

Session Classification: Artificial Intelligence

Track Classification: Track 10: Artificial Intelligence (AI)