Contribution ID: **6** Type: **Oral Presentation**

A study of foundation models for event classification in collider physics

Thursday, 20 March 2025 14:25 (25 minutes)

This study aims to improve the performance of event classification in collider physics by introducing a foundation model in deep learning. Event classification is a typical problem in collider physics, where the goal is to distinguish the signal events from the background events as much as possible to search for new phenomena in nature. Although deep learning can provide significant discrimination power in this event classification by exploiting its large parameter space, a large amount of data is necessary to maximize its performance. Because there are many data analyses that target various signal events, such as Higgs boson measurements and new particle searches, generating a large amount of training data using Monte Carlo simulations is computationally expensive. To address this problem, this presentation proposes a foundation model that can efficiently train the target even classification using a small amount of training data.

A foundation model is a pre-trained model, which is actively discussed in other fields such as natural language processing. A foundation model is usually trained on a large amount of unlabelled data, and then transferred and fine-tuned to downstream tasks, which involve common features, to efficiently train the model even with a small amount of data. In applying this foundation model concept to our field, the following novelties are introduced in this study. First, the real particle collision data collected by the CMS experiments are used to train the foundation model. Second, data augmentation techniques based on physics knowledge are applied. The advantage of using real data is that there is no need to generate a large amount of training data using simulations, which saves computing resources. However, due to the limited availability of true information in the real data, we have developed self-supervised learning techniques. Data augmentation also plays an important role because the amount of data needs to be increased as much as possible to build a robust foundation model. Based on our physics knowledge, the Lorentz transformations are applied to the events to increase the data patterns. Details of the self-supervised learning methods and data augmentations will be discussed in the presentation. The performance improvements of event classification by introducing the foundation model will also be shown.

Primary authors: TANAKA, Junichi (University of Tokyo); SAITO, Masahiko (ICEPP, The University of Tokyo); MORINAGA, Masahiro; KISHIMOTO, Tomoe (KEK)

Presenter: KISHIMOTO, Tomoe (KEK)

Session Classification: Artificial Intelligence (AI) - I

Track Classification: Track 10: Artificial Intelligence (AI)