T 119: Data, AI, Computing 8 (foundational & transformer models)

Time: Friday 9:00–10:15 Location: Geb. 30.33: MTI

T 119.1 Fri 9:00 Geb. 30.33: MTI

Finetuning Foundation Models for Joint Analysis Optimization — •Matthias Vigl, Nicole Hartman, and Lukas Heinrich — TUM

Most searches at the LHC employ an analysis pipeline consisting of various discrete components, each individually optimized and later combined to provide relevant features used to discriminate SM background from potential signal. These are typically high-level features constructed from particle four-momenta. However, the combination of individually optimized tasks doesn't guarantee an optimal performance on the final analysis objective. In this study, we show how an analysis would benefit from adopting an end-to-end ML optimization approach. Specifically, we investigate the impact of jointly optimizing particle identification and signal vs background discrimination exploiting the transformer-based ParT architecture [arXiv:2202.03772] as foundation model, showing the effectiveness of finetuning in the case of multi jets final states with CMS open data [DOI:10.7483/OPENDATA.CMS.JGJX.MS7Q].

T 119.2 Fri 9:15 Geb. 30.33: MTI

adaptive generative modeling for High-Granularity Calorimeters — •Lorenzo Valente — Institut für Experimentalphysik, University of Hamburg, Germany

Simulating particle colliders in their entirety presents a substantial computational challenge for researchers. Detector simulations are among the most resource-intensive phases of this process. Deep generative models could be a potential solution since they have already been proven to speed up simulations.

The growing volume of data from upcoming high-energy physics experiments, including higher collider luminosities and highly granular calorimeters, requires the development of artificial intelligence algorithms capable of combining knowledge across different domains. Unfortunately, conventional deep learning algorithms struggle with handling multiple datasets. Research in domain adaptation involves creating methodologies to bridge the divide between datasets, enabling the construction of models that exhibit high performance across diverse domains simultaneously.

In this contribution, we illustrate how a more universal domain adaptation approach, utilizing the transfer learning method, enhances the flexibility of the model. Specifically, we showcase its effectiveness in data generation for different calorimeter geometries.

T 119.3 Fri 9:30 Geb. 30.33: MTI

Leveraging Transformer Models for Gamma-Hadron Separation in SWGO — •Markus Pirke, Jonas Glombitza, Martin Schneider, and Christopher van Eldik for the SWGO-Collaboration — ECAP, FAU Erlangen-Nürnberg

The Southern Wide-field Gamma-ray Observatory (SWGO) is a proposed next-generation water-Cherenkov gamma-ray observatory in the Southern Hemisphere, thus being complementary to other water-Cherenkov detectors like HAWC (Mexico) and LHAASO (China), which are both located in the Northern Hemisphere. One of the primary challenges of the water-Cherenkov technique, is the effective discrimination of gamma-ray signals from the prevalent hadronic background. Several techniques have been developed in the past, primarily relying on human-designed discrimination variables.

In other scientific areas, recent advancements in deep learning have revealed that employing an end-to-end learning approach, which involves using raw data without the inclusion of handcrafted designed features, frequently improves the performance. One specific deep learning architecture is the Transformer. The self-attention mechanism of the Transformer, initially developed for tasks in natural language processing, offers a promising approach to efficiently handle the complex and variable-sized data in a ground-based observatory with high multiplicities. In this work, this approach will be investigated specifically for Gamma-Hadron separation in SWGO. Performance will be evaluated and additionally the inner workings, meaning the individual building blocks and their functions, of the Transformer will be explained.

T 119.4 Fri 9:45 Geb. 30.33: MTI

Point-Clouds based Diffusion Model on Hadronic Shower — •Martina Mozzanica[1], Erik Buhmann[1], Frank Gaede[2,3], Gregor Kasieczka[1,3], Anatolii Korol[2], William Korcari[1], Katja Krüger[2], and Peter McKeown[2] — [1]Institut für Experimentalphysik, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany — [2]Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany — [3]Center for Data and Computing in Natural Sciences CDCS, Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany

Simulating showers of particles in highly-granular detectors is a key frontier in the application of machine learning to particle physics. Achieving high accuracy and speed with generative machine learning models can enable them to augment traditional simulations and alleviate a major computing constraint.

Recent developments have shown how diffusion based generative shower simulation approach that do not rely on a fixed structure, but instead generates geometry-independent point clouds are very efficient. We present an extension to a point-cloud based diffusion model, i.e. CaloClouds, previously applied only to electromagnetic showers of the International Large Calorimeter (ILD).

The works focuses on the more challenging hadronic showers, namely pion showers, and introduces a more advanced architecture that successfully deals with the increasing complexity of the data, i.e. the attention mechanism.

T 119.5 Fri 10:00 Geb. 30.33: MTI

Photon Energy Reconstruction using Machine Learning at the Pierre Auger Observatory — •Daniel Rech — Karlsruhe Institute of Technology (IAP), Karlsruhe, Germany

An energy reconstruction for photon-induced air showers at ultra-high energies ($\geq 10^{18}$ eV) is presented for the surface detector of the Pierre Auger Observatory. Photon showers have a signature that differs from that of hadron-induced showers: the photon shower composition is almost exclusively electromagnetic and they show a steeper lateral distribution function as well as a larger depth of the shower maximum. In order to improve the resolution of the energy prediction, a reconstruction method based on ML is taken into consideration and compared to the classical hadron shower reconstruction applied to photon-induced extensive air showers. Due to the high success rate in other areas of machine learning, the encoder stack of the so-called transformer architecture is explored as an alternative to the more traditional approach of convolutional networks. So far, no photon events in the Pierre Auger dataset have been unequivocally identified as photons, but the advances in ML could play a key role in detecting them in the future.