Graph Neural Networks in HEP

AUTOGRAPH

Automatic Unified Training and Optimization for Graph Recognition and Analysis with Pipeline Handling

Greta Brianti

Introduction

A fully automated pipeline dedicated to **flavour tagging** exploiting **Graph Neural Network** (GNN)-based algorithms is presented. The pipeline is designed to allow the user **flexible and simple model training and validation**. In the jet-graph structure, the associated charged particle tracks correspond to nodes totally connected. The network inferences could be exploited to study the **network performance** through the discriminant distribution and ROC curve or/and to evaluate the **importance of the features** which allows to understand the model behavior and the physics behind the graph construction.

AUTOGRAPH pipeline

User interface

Configuration file

Dataset setting

- Input file
- Number of events
- Train dataset fraction
- Number of variables per node
- Number of global variables
- Number of nodes per graph

Network setting

- Graph layers type
- Number of layers
- Number of hidden nodes
- Pooling layer

Training setting

- Epochs
- Learning rate
- Batch size

Storing setting

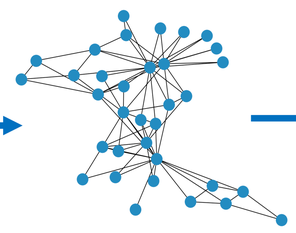
- Save performance
- Plot production
- Output directory

Automated steps

1. Pre-processing

Dataset setting

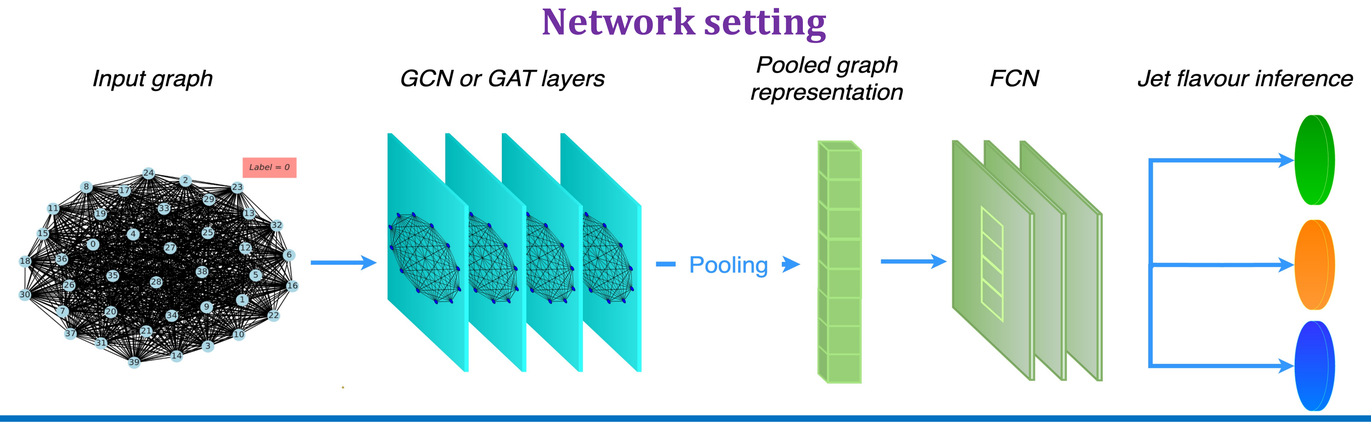
INPUT: n-tuples of variables in which each line collects the jet and the associated tracks features



OUTPUT: Training and validation datasets in Pytorch Geometric `DataLoader` format

2. Network architecture

Network setting



3. Train and validation

Training setting

INPUT: – Datasets, – Network

OUTPUT: Network inference

Performance evaluation
Loss function, discriminant

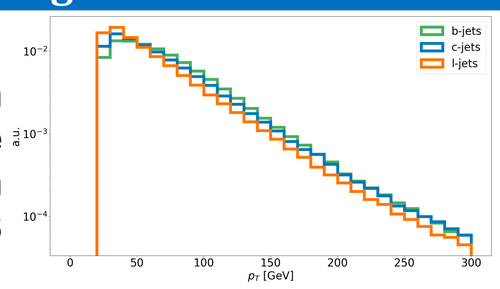
Feature ranking

Use case

Dataset simulation and pre-processing

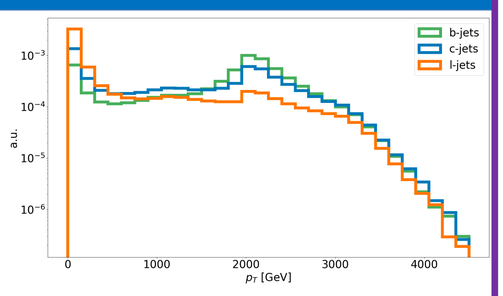
$t\bar{t}$ next-to-leading order (NLO)

The dataset production has required the integration of multiple frameworks. The response of particle detectors to the final-state particles has been produced with the fast simulation of the ATLAS detector.



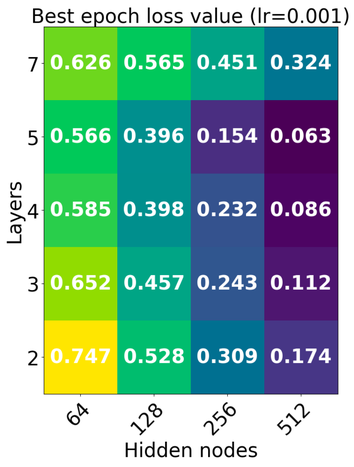
Z'H leading order (LO)

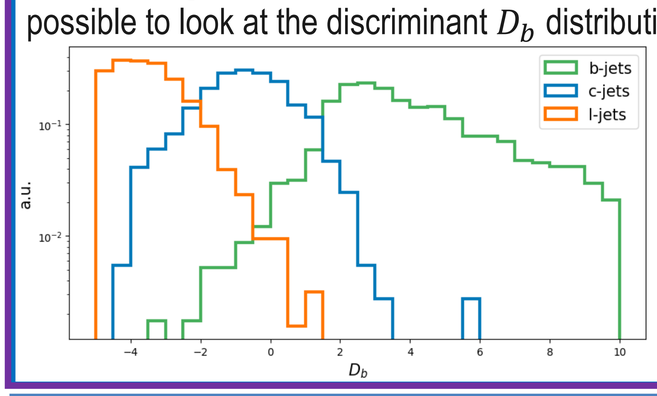
To extend the accessible transverse momentum range an additional dataset has been simulated with the fast simulation of the ATLAS detector. The Z' dark matter mediator candidate decay has been constrained in the hadronic channel.



Performance studies

To test the pipeline and study the network performances, I exploit 10'000 events from the **higher-p_T Z'H LO** dataset. I trained different network architectures by changing the Graph Convolutional Layer and hidden node numbers. The results are reported in the right image as loss value at the best epoch heat map. To evaluate the ability of the network in separating the contributions from the b-,c-, and l-jets it is possible to look at the discriminant D_b distribution [1].

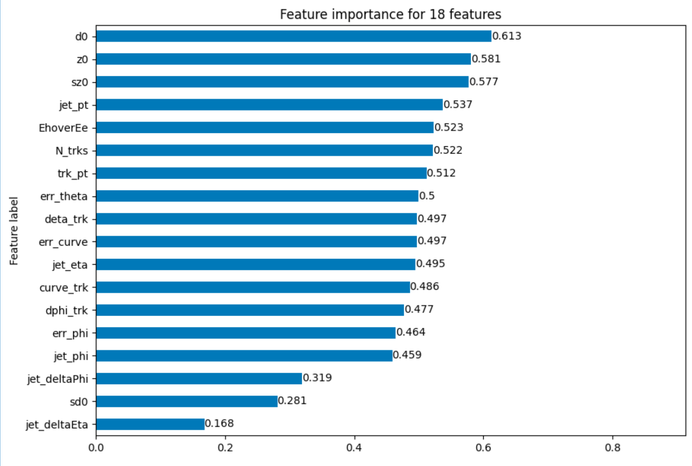




The figure shows the D_b distribution plot for the network corresponding to the minimum loss value. A good discrimination between flavours is reached with this architecture.

Features ranking

The feature ranking is directly performed from the pipeline with the Pytorch Geometric Explainer method [2]. The variable names and the number of features that we would like to classify are required. For this ranking, the network corresponding to the minimum loss value has been used.



18 features have been evaluated: 7 global and 11 per node.
The impact parameters, longitudinal d_0 and transverse z_0, are the most important variables for the network performance.
In fact, due to the B-hadron's lifetime, the b-jets have larger impact parameter values.

Conclusion

The pipeline presented is a tool for Graph Neural Networks (GNNs) training in the flavour tagging environment. The increasing use of GNNs needs a framework to accelerate the optimization and evaluation processes. The features ranking is a critical step for the network validation and to understand the physics beyond the graph-structured jets.

References

[1] *ATLAS sensitivity to Two-Higgs-Doublet models with an additional pseudoscalar exploiting four top quark signatures with 3ab−1 of √s = 14 TeV proton-proton collisions.* URL:https://cds.cern.ch/record/2645845.

[2] *Matthias Fey and Jan Eric Lenssen. "Fast Graph Representation Learning with PyTorch Geo-metric". In: CoRR abs/1903.02428 (2019). arXiv: 1903.02428*