Contribution ID: **86**   Type: **Poster**

ML-Based Top Taggers: Performance, Uncertainty and Impact of Tower & Tracker Data Integration

Tuesday, 12 December 2023 12:15 (1 hour)

Abstract : Machine learning algorithms have the capacity to discern intricate features directly from raw data. We demonstrated the performance of top taggers built upon three machine learning architectures: a BDT that uses jet-level variables (high-level features, HLF) as input, while a CNN (miniature version of ResNet) trained on the jet image, and a GNN (LorentzNet) trained on the particle cloud representation of a jet utilizing the 4-momentum (low-level features, LLF) of the jet constituents as input. We found significant performance enhancement for all three classes of classifiers when trained on combined data from calorimeter towers and tracker detectors. The high resolution of the tracking data not only improved the classifier performance in the high transverse momentum region, but the information about the distribution and composition of charged and neutral constituents of the fat jets and subjets helped identify the quark/gluon origin of sub-jets and hence enhances top tagging efficiency. The LLF-based classifiers, such as CNN and GNN, exhibit significantly better performance when compared to HLF-based classifiers like BDT, especially in the high transverse momentum region. Nevertheless, the LLF-based classifiers trained on constituents' 4-momentum data exhibit substantial dependency on the jet modeling within Monte Carlo generators. The composite classifiers, formed by stacking a BDT on top of a GNN/CNN, not only enhance the performance of LLF-based classifiers but also mitigate the uncertainties stemming from the showering and hadronization model of the event generator. We have conducted a comprehensive study on the influence of the fat jet's reconstruction and labeling procedure on the efficiency of the classifiers.

Presenter: SAHU, Rameswar (IOP, Bhubaneshwar)

Session Classification: Poster Session 1