Contribution ID: **35**   Type: **Hackathon project proposal**

Exploitation of jets features to tag VBF like events in high energy particle physics

Friday, 4 October 2024 10:40 (20 minutes)

The goal of the project is to develop a tagger for distinguishing VBF-like signal events from background ones using properties of the two leading jets (tag jets) and eventually the third one. The key idea is to exploit the color flow patterns in signal events, where the QCD activity between jets is greater between each jet and the beam spot with respect to the region between the two tag jets. Relevant features for the tagger may include jet kinematic properties, jet substructure variables, and color flow observables. Signal (VBF) and background (Drell Yan) samples will be provided for training and validation, the study will be done at parton level.

Project proposal: general context

The Large Hadron Collider (LHC) at CERN is the world's most powerful particle accelerator, designed to collide protons (and heavy ions) at extremely high energies, reaching up to 13.6 TeV and operating at a collision rate of 40 MHz. However, most of these proton-proton collisions result in low-energy "soft" interactions, which do not produce particles of interest for high-energy physics studies. To manage the high event rate, the LHC experiments employ a trigger system that selects and saves only the most promising, "hard scattering" events for further analysis.

Among the four main experiments at the LHC, ATLAS and CMS have a broad physics program aimed at exploring both the Standard Model (SM) and potential new physics. One process of particular interest to study the electroweak sector of the standard model is Vector Boson Fusion (VBF).

In VBF a quark from each of the incoming LHC protons radiates off a heavy vector boson. These bosons interact to produce another boson. The initial quarks that first radiated the vector bosons are deflected only slightly and travel roughly along their initial directions, producing two forward jets (tag jets) with minimal activity between them

Project proposal: description of the problem

While the trigger system of the experiments effectively filters out many soft events, background processes that mimic the signal of interest can still pass through. In the case of VBF, distinguishing this signal from background events like Drell-Yan remains challenging. In a Drell-Yan process, a quark of one proton and an antiquark of another proton annihilate, creating a virtual photon or Z boson which then decays into a pair of oppositely charged leptons.

Both processes can produce similar final-state particles, making it difficult to differentiate between signal and background. Therefore, it is necessary to develop a more sophisticated classifier that can accurately discriminate VBF events from background events based on jet kinematic properties, substructure, and color flow patterns.

Machine learning methods

The task is classification/tagging. A classifier would work, but a different approach with a CNN that exploits the tag jets images could be interesting to explore.

Input dataset

the notebooks provided here:
https://github.com/AuroraPerego/VBS_tagger/tree/tagger/notebooks
can be used to read and plot the data
two files, one for the signal (70MB) and one for the background (80 MB), each of them containing 10k events,
saved ad ROOT TTrees.

Goal and FOM

the target can be the ROC and the AUC, the goal is to have high purity for the tagging.

Primary author: PEREGO, Aurora (Istituto Nazionale di Fisica Nucleare)

Session Classification: Hackathon