EuCAIFCon 2024

EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE

Amsterdam, 30 April - 3 May

Anomaly detection search for BSM physics in ATLAS experiment at LHC

Francesco Cirotto[1,3], Francesco Conventi[2,3], Elvira Rossi[1,3]

[1]Università degli studi di Napoli Federico II
[2]Università degli studi di Napoli Parthenope
[3]INFN Napoli

WHERE IS LHC GOING?

Finding anomalies in Jets

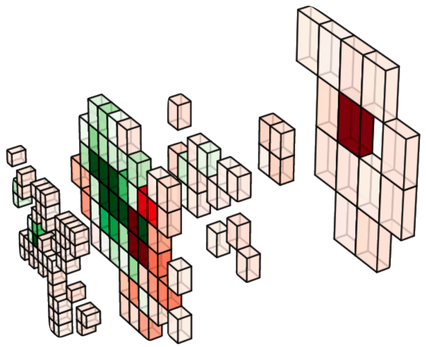
- No Physics Beyond Standard Model (BSM) has been observed at the LHC (yet!)

- The currently most used search paradigm is using model-dependent approaches
  - ↪What if these models have blind spots for unconventional new physics signatures?
  - ↪If there's new Physics in the current LHC data we can't miss it!

- Anomaly detection can find deviations in Standard Model events, without any signal dependency

Use jets as tools!

- Searches in full hadronic final states
  - ↪Investigate its substructure by studying jet constituents
  - ↪What we use: p_T, η, ϕ of each constituent



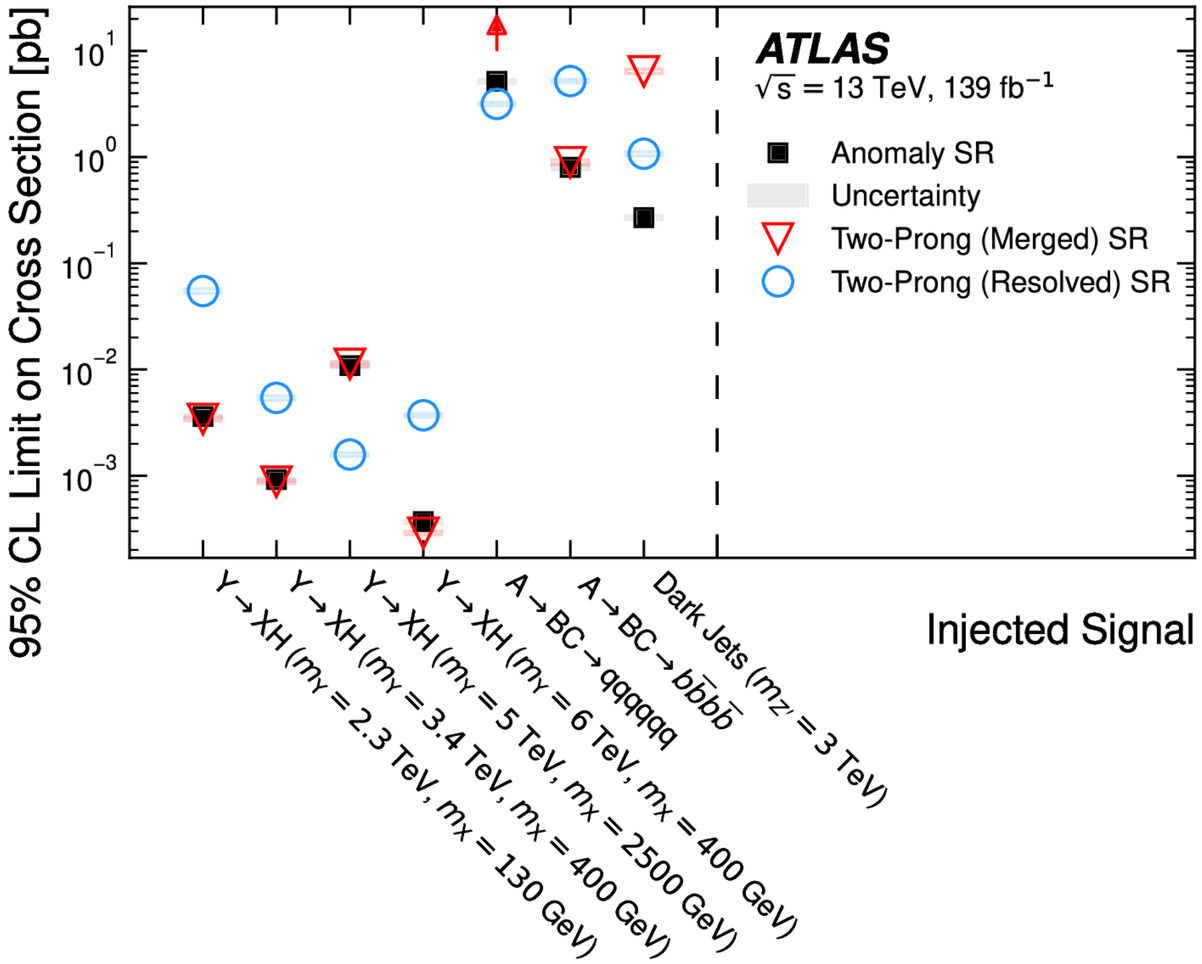
$Y \rightarrow XH$

The First Use of Unsupervised Learning on ATLAS Data

- The Y → XH analysis searches for heavy resonances decaying into a Higgs boson and new particle X in a fully hadronic final state

- Developed an unsupervised Variational Recurrent Neural Network to define an "Anomaly Signal Region)
  - ↪ Recurrent neural network that updates a VAE latent space at each time step, accommodating variable-length input sequences

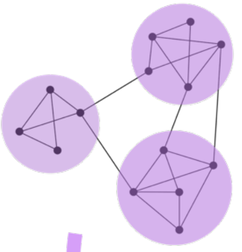
Can we improve the "classical" approach?

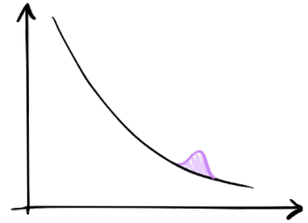


DIGGING DEEPLY WITH GRAPHS

A developing approach

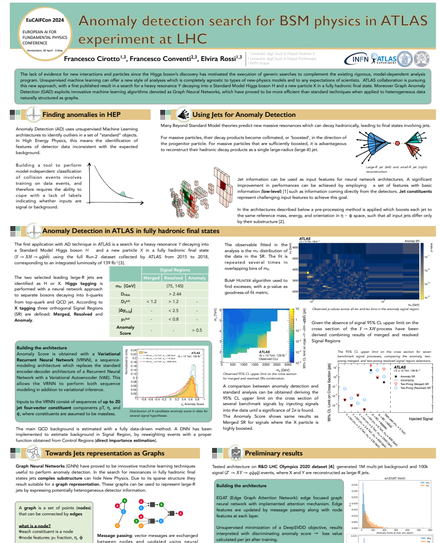
- Graph-structured data are ubiquitous across science, engineering, and many other domains
  - ↪Used to describe and analyze relations and interactions
  - ↪Can encapsulate object or event information





- Our strategy: to represent jets as graphs and then apply machine learning to build an anomaly detection algorithm
  - ↪Developing both **object and event level graphs** to detect anomalies

- Preliminary results on LHC Olympics dataset


EuCAPCon 2024
EUROPEAN IA FOR FUNDAMENTAL PHYSICS CONFERENCE
Anomaly detection search for BSM physics in ATLAS experiment at LHC
Francesco Cirotto[1,3], Francesco Conventi[2,3], Elvira Rossi[1,3]
[1] Università degli Studi di Napoli Federico II
[2] Università degli Studi di Napoli Parthenope
[3] INFN Sezione di Napoli
The lack of evidence for new interactions and particles since the Higgs boson's discovery has motivated the execution of generic searches to complement the existing rigorous, model-dependent analysis program. Unsupervised machine learning can offer a new style of analyses which is completely agnostic to types of new-physics models and to any expectations of scientists. ATLAS collaboration is pursuing this new approach, with a first published result in a search for a heavy resonance Y decaying into a Standard Model Higgs boson H and a new particle X in a fully hadronic final state. Moreover Graph Anomaly Detection (GAD) exploits innovative machine learning algorithms denoted as Graph Neural Networks, which have proved to be more efficient than standard techniques when applied to heterogeneous data naturally structured as graphs.
Finding anomalies in HEP
Using Jets for Anomaly Detection
Anomaly Detection (AD) uses unsupervised Machine Learning architectures to identify outliers in a set of "standard" objects. In High Energy Physics, this means the identification of features of detector data inconsistent with the expected background.
Building a tool to perform model-independent classification of collision events involves training on data events, and therefore requires the ability to cope with a lack of labels indicating whether inputs are signal or background.
Many Beyond Standard Model theories predict new massive resonances which can decay hadronically, leading to final states involving jets.
For massive particles, their decay products become collimated, or 'boosted', in the direction of the progenitor particle. For massive particles that are sufficiently boosted, it is advantageous to reconstruct their hadronic decay products as a single large-radius (large-R) jet.
Jet information can be used as input features for neural network architectures. A significant improvement in performances can be achieved by employing: a set of features with basic information (low-level) [1] such as information coming directly from the detectors, jet constituents represent challenging input features to achieve this goal.
In the architectures described below a pre-processing method is applied which boosts each jet to the same reference mass, energy, and orientation in η − φ space, such that all input jets differ only by their substructure [2].
Anomaly Detection in ATLAS in fully hadronic final states
The first application with AD technique in ATLAS is a search for a heavy resonance Y decaying into a Standard Model Higgs boson H and a new particle X in a fully hadronic final state ($Y \to XH \to qqbb$) using one full Run-2 dataset collected by ATLAS from 2015 to 2018, corresponding to an integrated luminosity of 139 fb^{-1} [3].
The two selected leading large-R jets are identified as H or X. Higgs tagging is performed with a neural network approach to separate bosons decaying into b-quarks from top-quark and QCD jet. According to X tagging three orthogonal Signal Regions (SR) are defined, Merged, Resolved and Anomaly.
Signal Regions
Variable | Merged | Resolved | Anomaly
m_Y [GeV] | (75, 145) | - | -
D_{bb} | > 2.44 | - | -
$D_2^{(\beta=1)}$ | < 1.2 | > 1.2 | -
ΔR_{bb} | - | < 2.5 | -
p_T^{jet} | - | < 0.8 | -
Anomaly Score | - | - | > 0.5
The observable fitted in the analysis is the m_Y distribution of the X and H candidates. It is reported several times in overlapping bins of m_Y.
Boost HUNTER algorithm used to find excesses, with a p-value as goodness-of-fit metric.
Given the absence of signal 95% CL upper limit on the cross section of the Y → XH process have been derived combining results of merged and resolved Signal Regions.
Building the architecture
Anomaly Score is obtained with a Variational Recurrent Neural Network (VRNN), a sequence modeling architecture which replaces the standard encoder-decoder architecture of a Recurrent Neural Network with a Variational Autoencoder (VAE). This allows the VRNN to perform both sequence modeling in addition to variational inference.
Inputs to the VRNN consist of sequences of up to 20 jet four-vector constituent components p_T^{rel}, η_i and ϕ_i where constituents are assumed to be massless.
The main QCD background is estimated with a fully data-driven method. A DNN has been implemented to estimate background in Signal Region, by reweighting events with a proper function obtained from Control Regions (direct importance estimation).
Towards Jets representation as Graphs
Graph Neural Networks (GNN) have proved to be innovative machine learning techniques useful to perform anomaly detection, in the search for resonances in fully hadronic final states jets complex substructure can hide New Physics. Due to its sparse structure they result suitable for a graph representation. These graphs can be used to represent large-R jets by expressing potentially heterogeneous detector information.
A graph is a set of points (nodes) that can be connected by edges
what is a node?
#each constituent is a node
#node features: p_T fraction, η, ϕ
when are nodes connected?
#Constituent distance-based edge feature
Message passing: vector messages are exchanged between nodes and updated using neural networks. At each iteration, every node aggregates information from its local neighborhood. At the end each node embedding contains more and more information from further reaches of the graph.
Preliminary results
Tested architecture on R&D LHC Olympics 2020 dataset [4] generated 1M multi-jet background and 100k signal ($Z' \to X Y \to qqbb$) events, where X and Y are reconstructed as large-R jets.
Building the architecture
EGAT (Edge Graph Attention Network): edge-focused graph neural network with implemented attention mechanism. Edge features are updated by message passing along with node features at each layer.
Unsupervised minimization of a DeepSVDD objective, results interpreted with discriminating anomaly score → loss value calculated per jet after training.
Event level score obtained by combining the 2 jets anomaly scores.
Very promising results with LHC Olympics dataset, AUC ~0.82 in event-level anomaly score approach.
The architecture will be tested on official ATLAS datasets.
References
[1] Kasieczka, G., Nachman, B. & Shih, D. Eur. Phys. J. C (2021) 81: 214. https://doi.org/10.1140/epjc/s10052-021-08791-4
[2] Dreyer, F. A., Salam, G. P. & Soyez, G. J. High Energ. Phys. (2020) 2020: 93. https://doi.org/10.1007/JHEP12(2020)093
[3] ATLAS Collaboration, Anomaly detection search for new resonances decaying into a Higgs boson and a generic new particle X in hadronic final states using $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector, ATLAS-CONF-2023-042
[4] The LHC Olympics 2020: a community challenge for anomaly detection in high energy physics, Rep. Prog. Phys. 84 (2021) 124201

Poster 76