LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS
partículas e tecnologia

Deep Learning Models in searches for new physics at colliders

Lomonosov Conference
25th August 2021

Miguel Crispim Romão
mcromao@lip.pt

One of the main goals of the LHC is to look for New Physics

1. Choose BSM signal you are looking for
2. Study favourable kinematic region and final state topology
3. Collect the data in such regime
4. Perform statistical tests on the data on the hypothesis of BSM being present
5. Profit (eventually)

Searches for New Physics at Modern Colliders

The Workflow Challenges

- An event is characterised by a collection of kinematic variables (jet and lepton masses, pT, eta, phi, multiplicities, b-tags, etc) => Multivariate Analysis
    - What are the best discriminating variables? => Use a single Machine Learning discriminator! (Neural Network or Gradient Boosted Trees)
    - What if the signal region on these variables change as we change the parameters of New Physics?
- An explicit New Physics hypothesis is tested
    - What if another New Physics signal is presented instead?
    - What if we are forgetting to consider a realised New Physics case (for example something more exotic that is not covered in standard analyses)

Transferability of Deep Learning Models in Searches for New Physics at Colliders

MCR, N. F. Castro, R. Pedro, T. Vale

Phys.Rev.D 101 (2020) 3, 035042 [1912.04220]

- How does an NN classifier, trained to separate a specific signal from background, behave when shown a new signal?
- How does this impact upper limits on New Physics?
- Focused on three classes of signals:
  - FCNC
  - VLQ from SM production
  - VLQ from Heavy Gluon production

Transferability of Deep Learning Models
Analogy

Jungle is the Background (SM events) and we want to find monkeys (a BSM candidate)

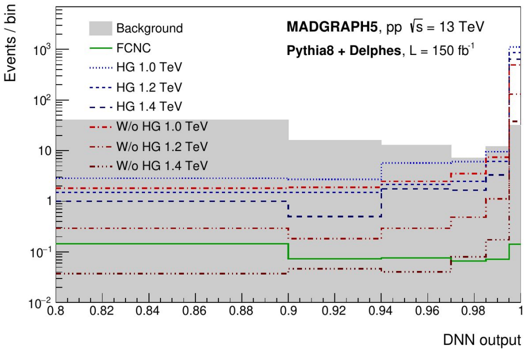
What happens if instead of monkeys there is another animal in the data?

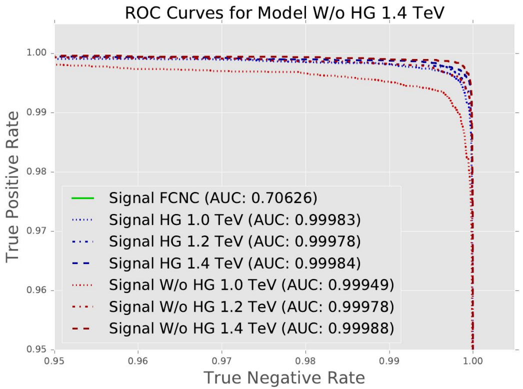
Would an NN still find the signal?

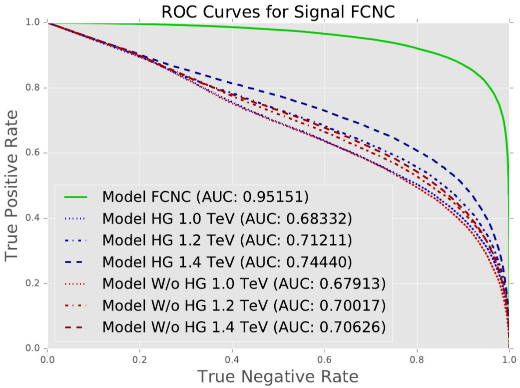
Transferability of Deep Learning Models
Methodology

- For each signal train a supervised DNN classifier
- Use each trained DNN to predict on every combination signal-background
- Assess how discrimination deteriorates as we present a different signal to each DNN through upper limits on expected cross-section

Transferability of Deep Learning Models

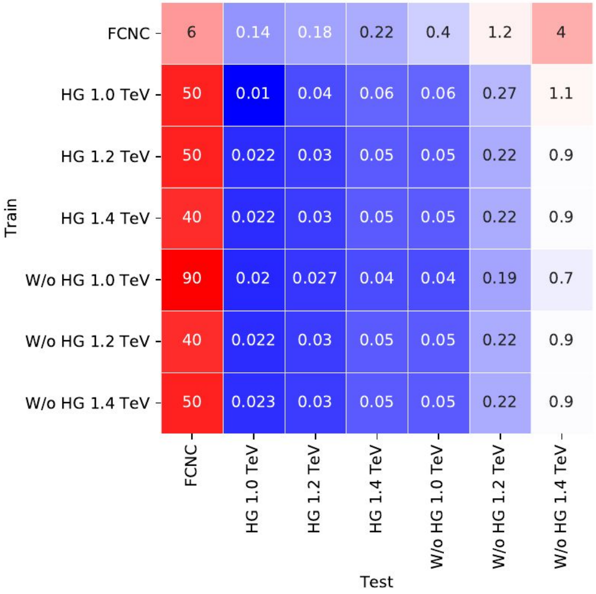


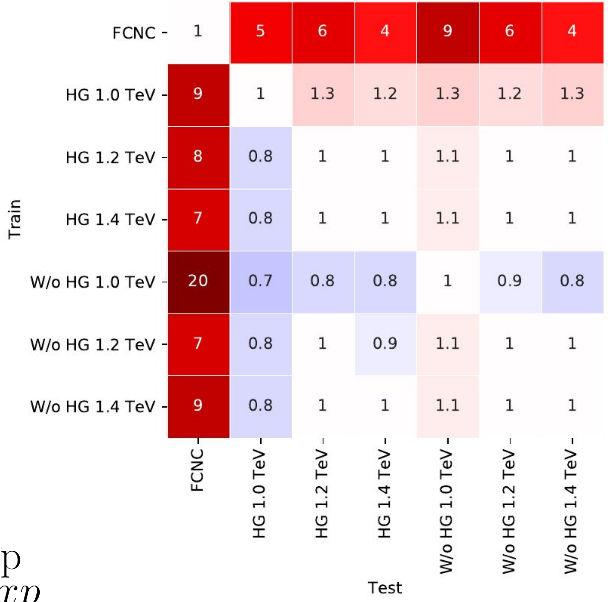




Transferability of Deep Learning Models

Upper Limits





$$\mu = \frac{\sigma_{exp}^{\mathrm{up}}}{\sigma_{th}}$$

Could we not just focus on the jungle?

Since we don't know what BSM candidate is realised in nature, it seems it would be better if we could develop a way of identifying **any type of non SM phenomena**

Unsupervised Methods for New Physics Searches

- Growing interest in Unsupervised approaches to isolate New Physics from SM Background
- Anomaly Detection ML algorithms are finding their way into HEP to help this out
  - 1805.02664, 1808.08992, 1811.10276, 1902.02634, 1903.02032, ...
- A comprehensive live review of ML in HEP curated by CERN's IML WorkGroup: https://github.com/iml-wg/HEPML-LivingReview

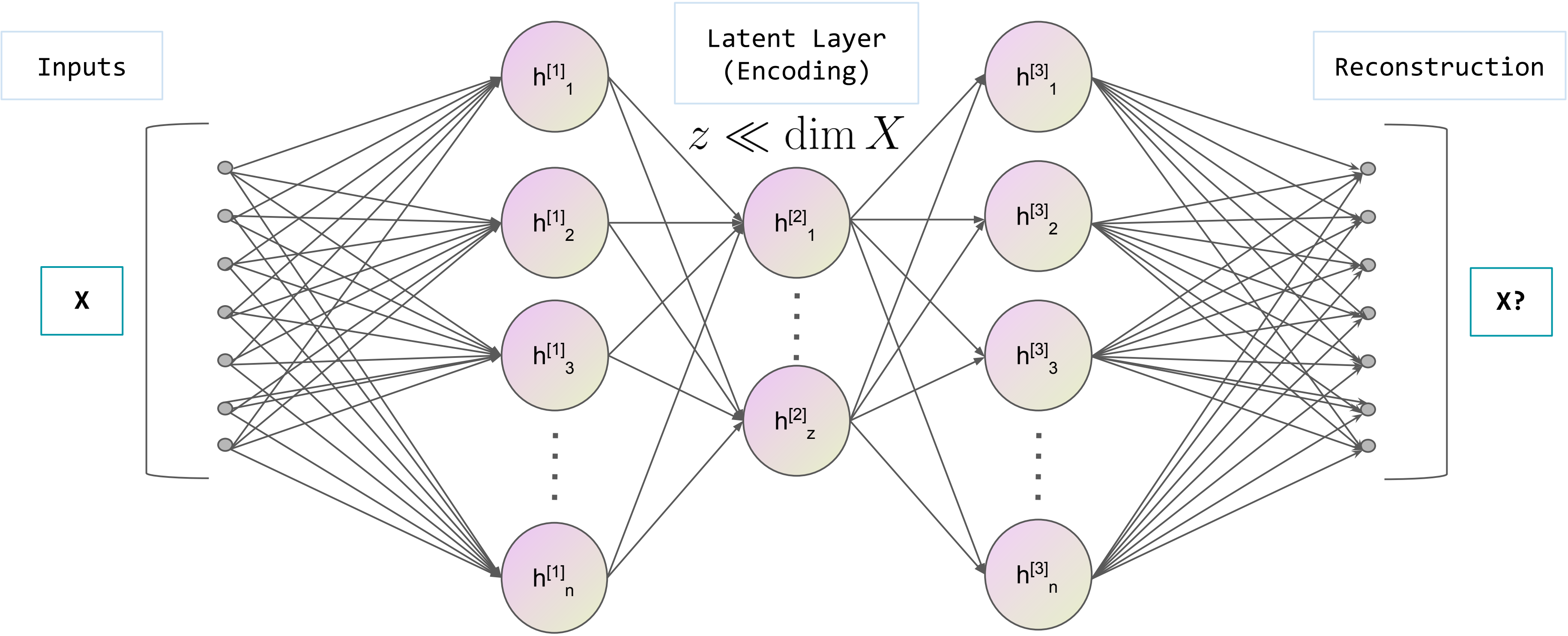
Finding New Physics without learning about it: Anomaly Detection as a tool for Searches at Colliders

MCR, N. F. Castro, R. Pedro

Eur.Phys.J.C 81 (2021) 1, 27
[2006.05432]

- We kept the same signals
  - FCNC
  - VLQ from SM production
  - VLQ from Heavy Gluon production
- We compared four AD algorithms
  - Auto-Encoder
  - Deep-SVDD
  - Isolation Forest
  - Histogram Based

Finding New Physics without learning about it
Auto-Encoder



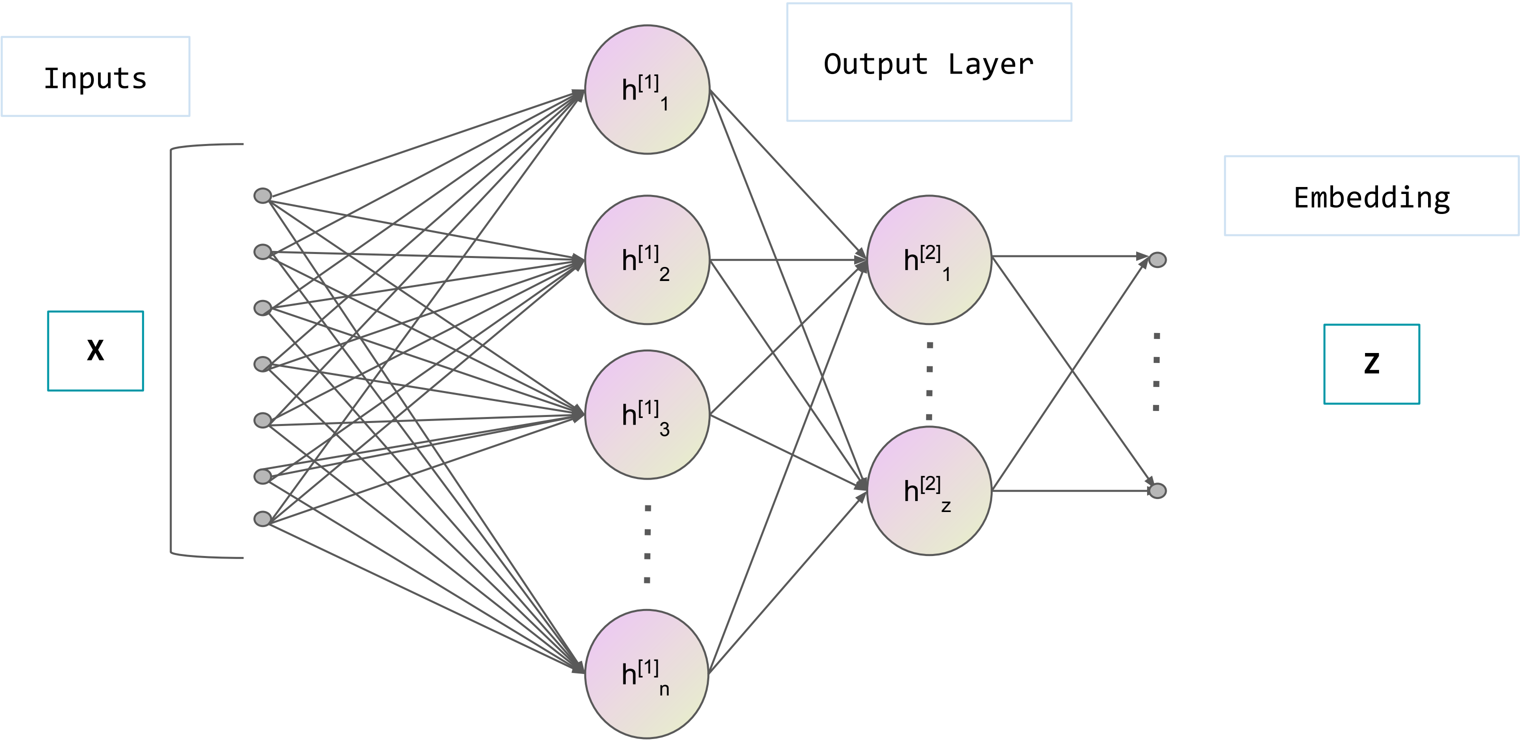
Finding New Physics without learning about it
Auto-Encoder

- The Network is trained by minimising the reconstruction error

$$L = \frac{1}{N}\sum_{i=1}^{N} |x_i - \mathrm{AE}(x_i)|^2$$

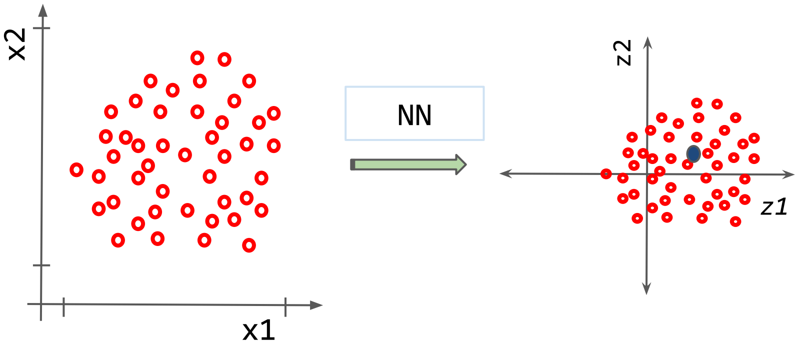
- In principle, events that are easier to reconstruct are the most common (or at least carry the most common relations between the variables)
- Reconstruction error of an event can be a measure of how rare (anomalous) it is => BSM events should have higher reconstruction error

Finding New Physics without learning about it
Deep-SVDD



Finding New Physics without learning about it
Deep-SVDD

- Before any training, the NN is just a map from the input space to some embedding space



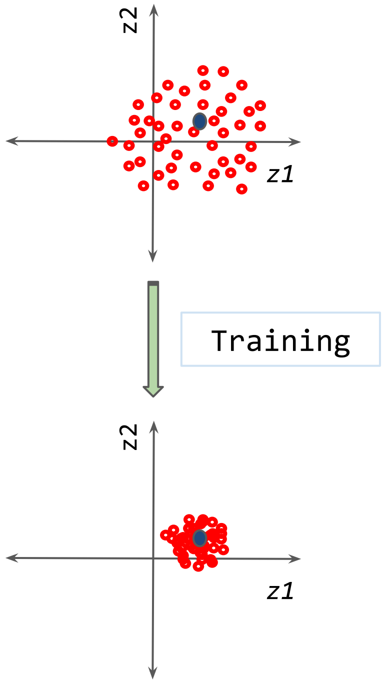
- In this space we can find a “centre of mass”, c, of the points

Finding New Physics without learning about it
Deep-SVDD

- The Network is trained by minimising the distance to the centre of mass

$$L = \frac{1}{N}\sum_{i=1}^{N} |c - \mathrm{NN}(x_i)|^2$$

- The bulk of the distribution will be easier to bring to the centre, the rarer events will be further away
- The distance to c becomes then a natural interpretation for *outlyingness* of an event

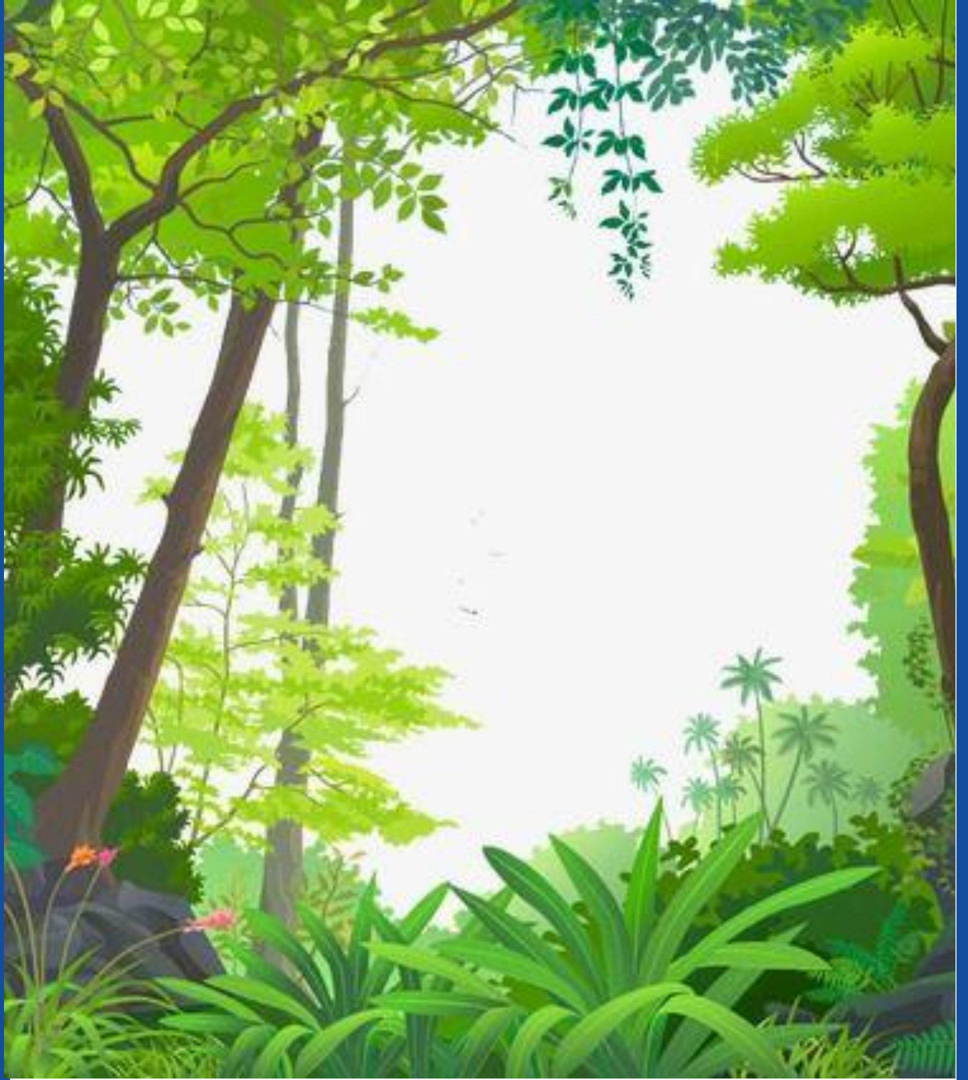


Train only on Standard Model

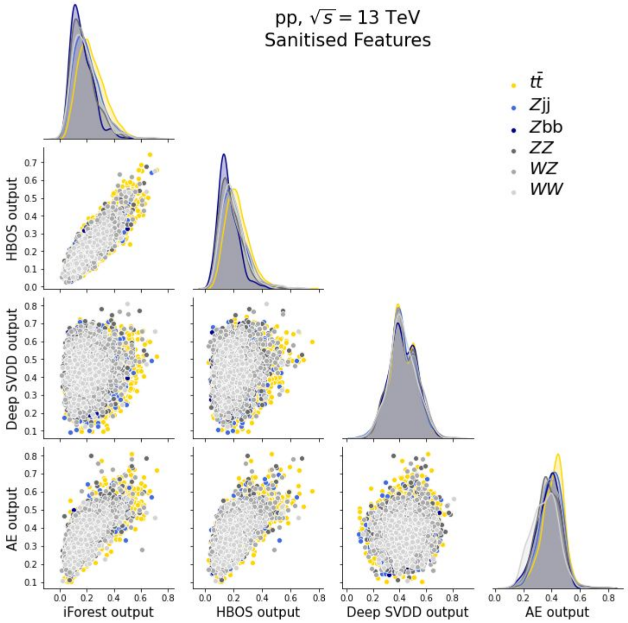
This way we are learning what a jungle looks like and hopefully we will be able to find any animal!

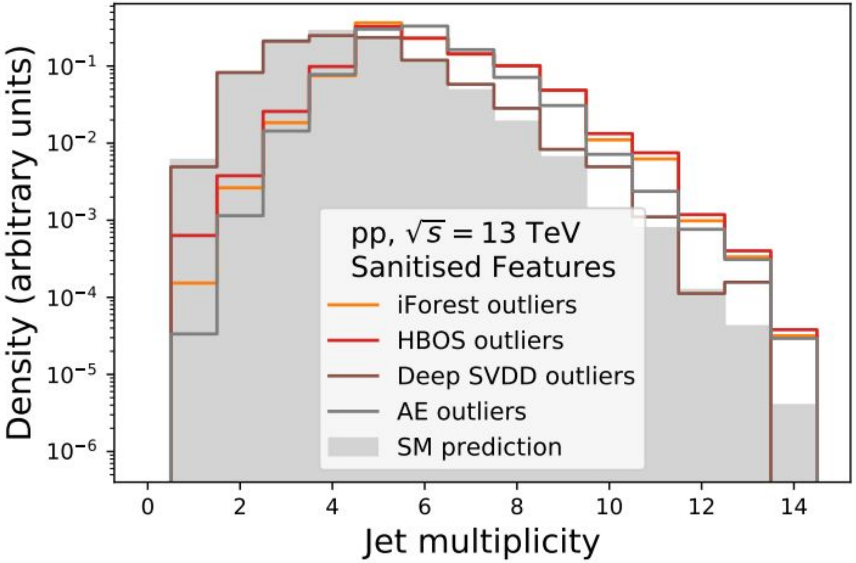
Are different algorithms correlated?

Are they focusing on the same characteristics?



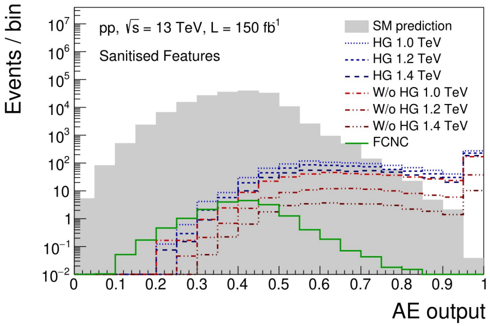
Finding New Physics without learning about it
Results 1: Are all AD algorithms created equally?

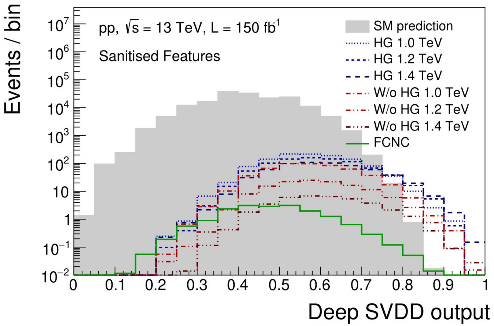


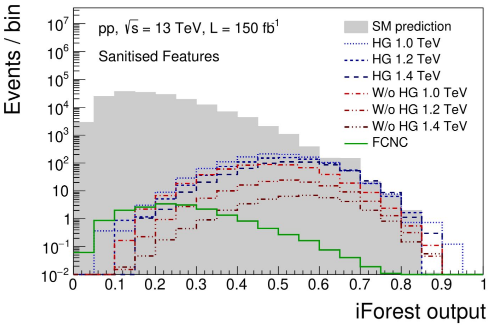


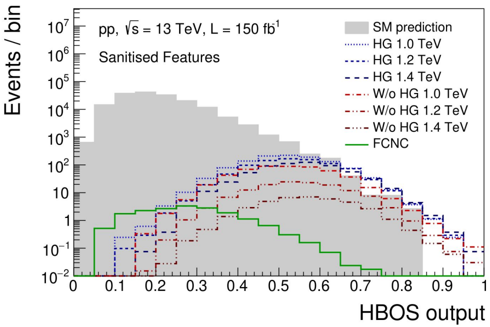
Finding New Physics without learning about it

Results 2: Can they find animals?

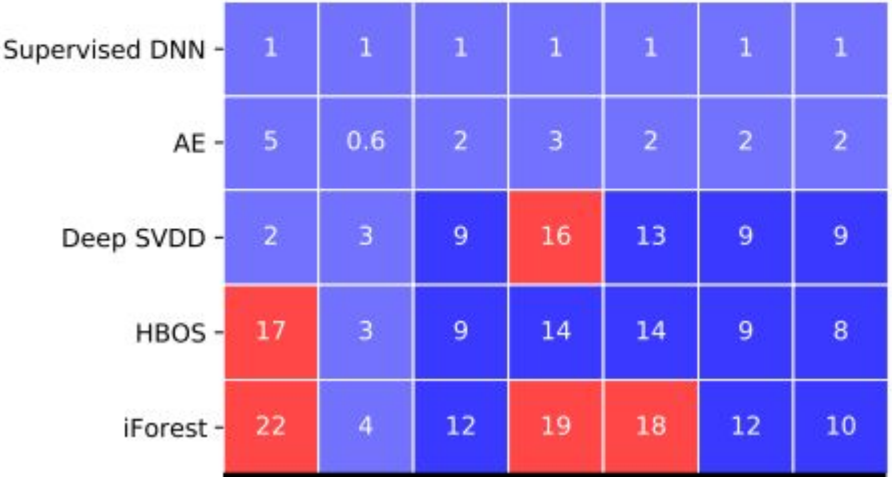








Finding New Physics without learning about it
Results 3: Can we search for New Physics?



Ongoing and future work

Ongoing and Future Work

- Systematic study on how we can maximise the sensitivity of Anomaly Detection methods
- A sensitive analysis for a BSM model which will include an unsupervised analysis alongside the supervised one
- Still a lot to understand on how and when each Anomaly Detection methods works best for different classes of BSM signals and topologies
- Need a robust “blind” statistical test to use on the new discriminat provided by the Anomaly Detection methods

Conclusions

Conclusions

- NN provide very versatile solutions for generic searches
  - Supervised NN classifiers are able to find other signals
  - Unsupervised architectures provide at most an order of magnitude of degradation in sensitivity against supervised
- Unsupervised methods are getting a lot of attention and interest in the community and can provide a BSM independent solution to search for NP
- Ongoing and future work:
  - Motivate phenomenologist to include unsupervised sensitivity analyses of their models
  - Systematic studies on how to confidently deploy these methods in production/experiment

Thanks!

mcromao@lip.pt

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824093."

POCI/01-0145-FEDER-029147
PTDC/FIS-PAR/29147/2017

Fundação para a Ciência e a Tecnologia

UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional

n+1.
Backups

Transferability of Deep Learning Models
The Background

Now available on Zenodo!
https://zenodo.org/record/5126747

- A SM cocktail sample was produced in MadGraph5+Pythia8+Delphes
  - 8M Z+J, 3M ttbar, 1.5M per diboson sample
- Targeted processes with dilepton final state, at least one b, and HT > 500 GeV
- To guarantee statistics at the tails of the distributions we applied event filter at parton level in pT slices
- The events are represented by variables from the reconstructed objects:
  - (η, φ, pT , m) for 5 leading jets and large-radius jets
  - N–subjetiness of the leading large-radius jet (τ_N with N = 1, 2, .., 5)
  - (η, φ, pT) of the 2 leading electrons and muons
  - Multiplicites, (ET , φ) of the missing transverse energy (MET)

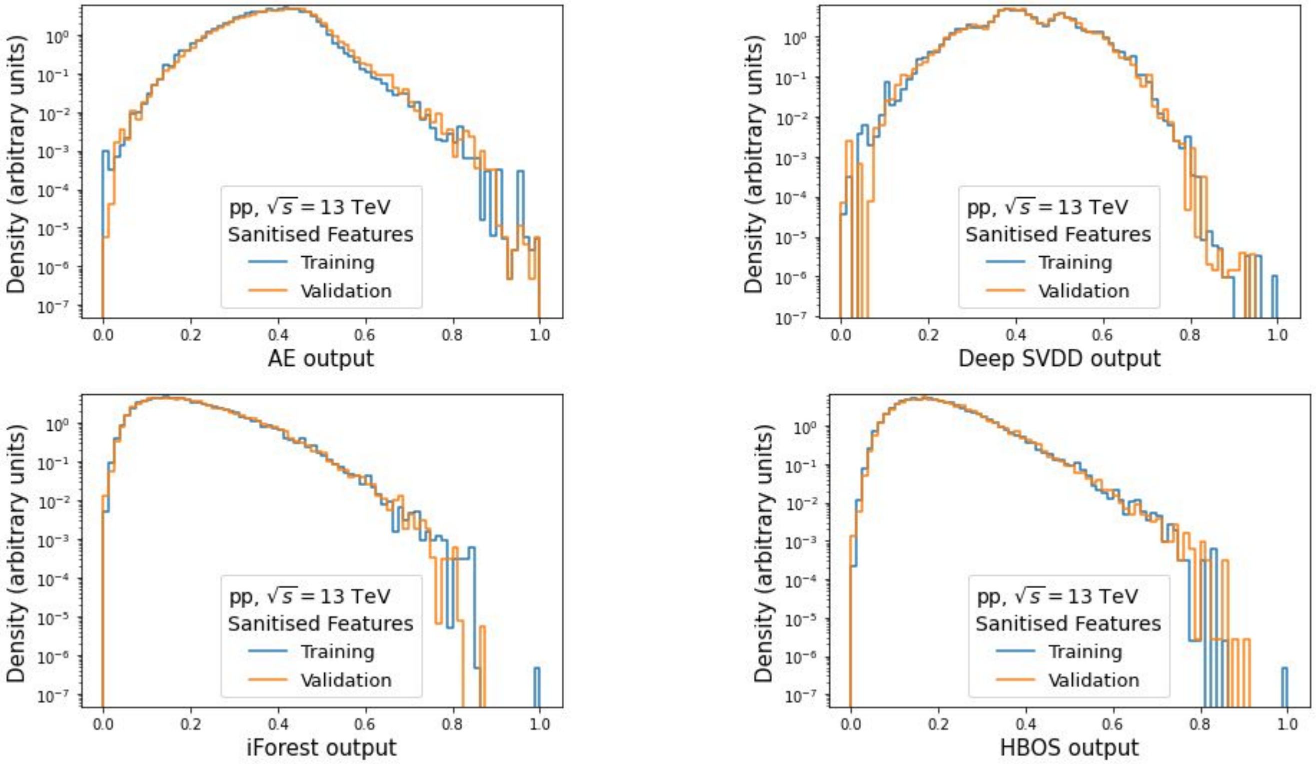
Transferability of Deep Learning Models

The Signals

Now available on Zenodo!
https://zenodo.org/record/5126747

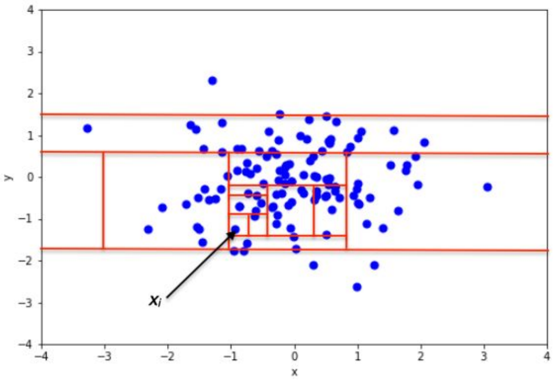
- 7 samples of BSM signals over three classes
- FCNC interaction in single top-quark production
- Vector-Like T quarks produced via SM gluon with three different masses
    - 1.0 TeV
    - 1.2 TeV
    - 1.4 TeV
- Vector-Like T quarks produced via BSM heavy (3TeV) gluon with three different masses
    - 1.0 TeV
    - 1.2 TeV
    - 1.4 TeV

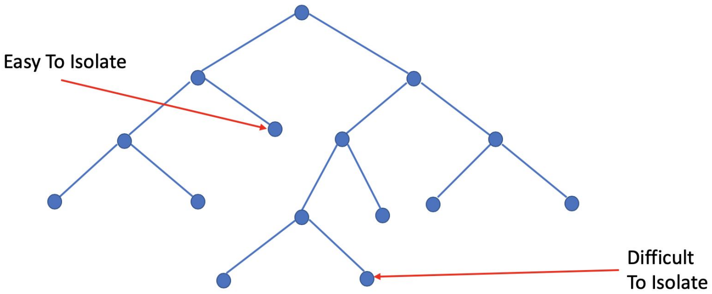
Finding New Physics without learning about it
Results 0: When they see new jungle



Finding New Physics without learning about it
Isolation Forest

- Recursively partition the data with random cuts
- These cuts can be represented as a tree
- Rare events will be easier to isolate
- Anomaly score given by the inverse of how many nodes it took to isolate

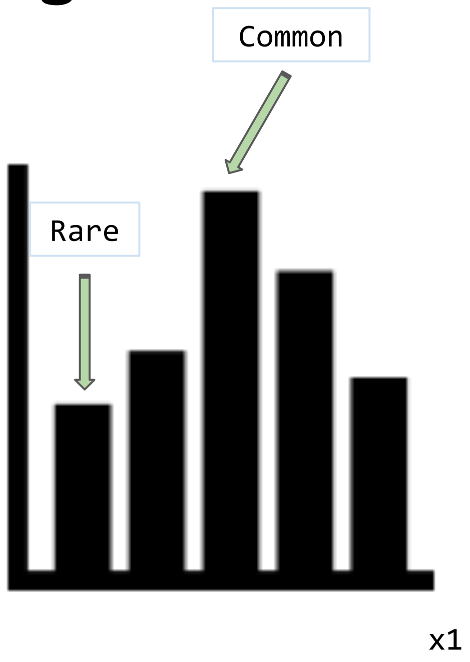




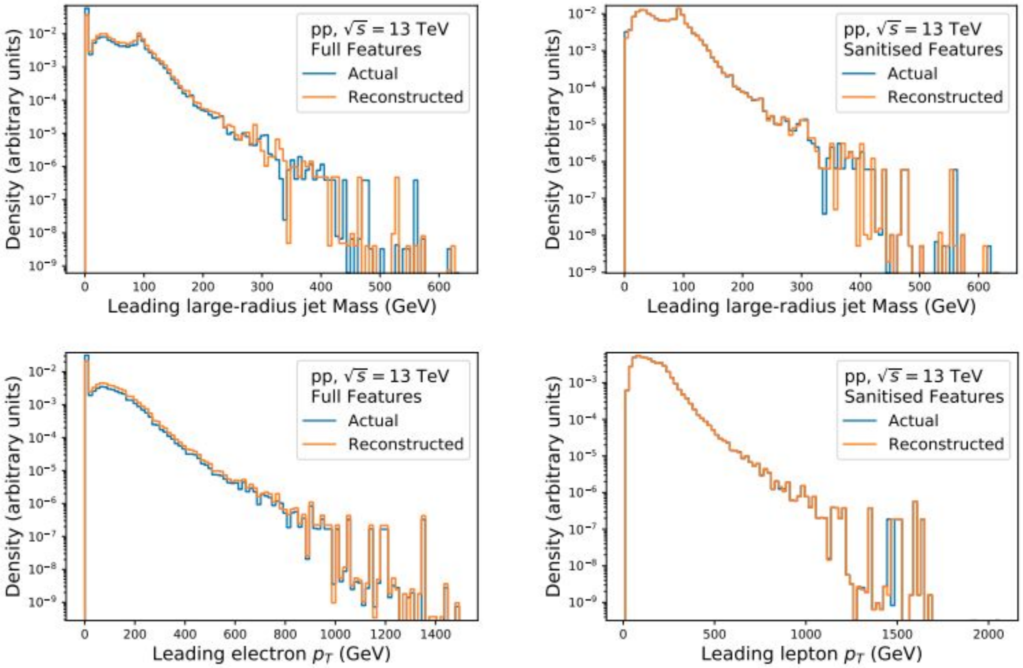
Finding New Physics without learning about it

Histogram Based

- Compute histograms for all variables
- Rare events will more often be in bins of smaller height
- Anomaly score given by the sum of the Log of the heights of each bin an event occupies

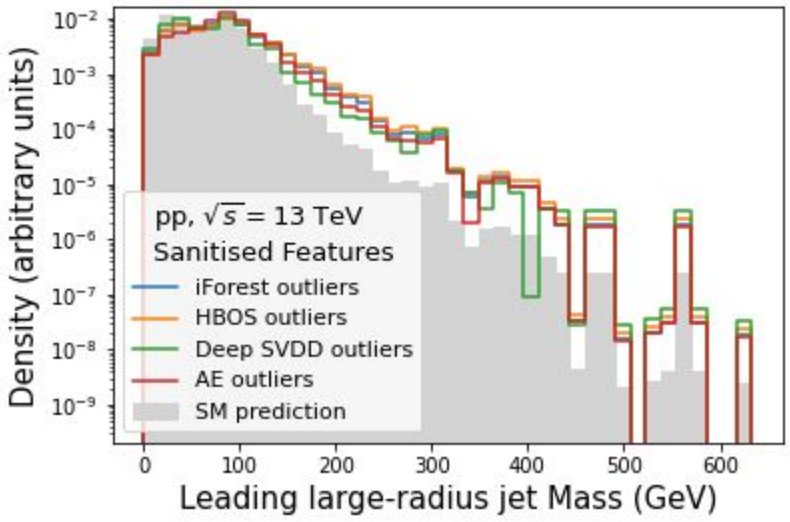


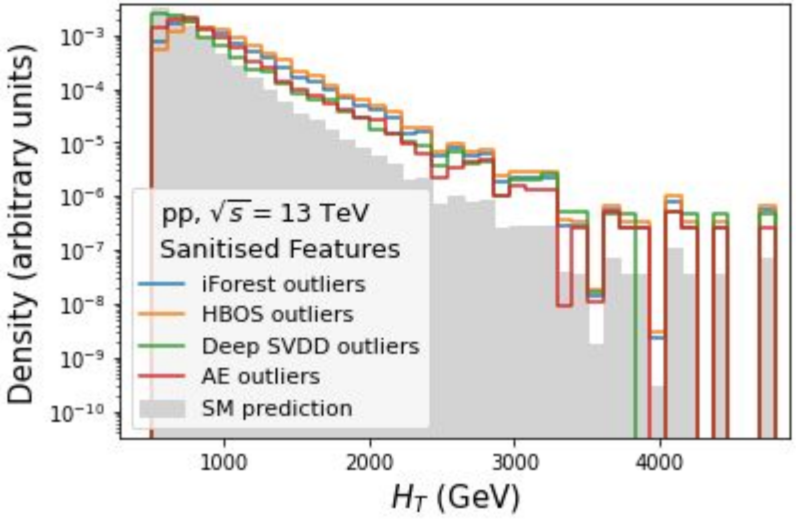
Finding New Physics without learning about it
Sanitised Features



Backups

AD outliers are data outliers





Backups

AD mus

Model	Benchmark Signal						
	FCNC	HG			No HG		
		1.0 TeV	1.2 TeV	1.4 TeV	1.0 TeV	1.2 TeV	1.4 TeV
			Full features				
Supervised DNN	6^{+3}_{-2}	$0.011^{+0.007}_{-0.004}$	$0.015^{+0.008}_{-0.005}$	$0.016^{+0.009}_{-0.005}$	$0.03^{+0.02}_{-0.01}$	$0.08^{+0.04}_{-0.03}$	$0.20^{+0.12}_{-0.07}$
H_T	110^{+40}_{-30}	$0.14^{+0.07}_{-0.05}$	$0.16^{+0.08}_{-0.06}$	$0.16^{+0.08}_{-0.05}$	$0.4^{+0.2}_{-0.1}$	$1.0^{+0.5}_{-0.3}$	$1.8^{+0.9}_{-0.6}$
Deep SVDD	60^{+30}_{-20}	$0.29^{+0.14}_{-0.09}$	$0.32^{+0.15}_{-0.10}$	$0.4^{+0.2}_{-0.1}$	$0.8^{+0.4}_{-0.2}$	$1.9^{+0.9}_{-0.6}$	5^{+2}_{-1}
AE	30^{+10}_{-10}	$0.06^{+0.04}_{-0.02}$	$0.06^{+0.05}_{-0.02}$	$0.06^{+0.04}_{-0.02}$	$0.12^{+0.08}_{-0.04}$	$0.4^{+0.2}_{-0.1}$	$1.0^{+0.6}_{-0.3}$
HBOS	100^{+40}_{-30}	$0.15^{+0.07}_{-0.05}$	$0.17^{+0.08}_{-0.05}$	$0.19^{+0.09}_{-0.06}$	$0.4^{+0.2}_{-0.1}$	$1.0^{+0.5}_{-0.3}$	$2.7^{+1.2}_{-0.9}$
iForest	200^{+60}_{-40}	$0.22^{+0.11}_{-0.07}$	$0.26^{+0.13}_{-0.09}$	$0.3^{+0.2}_{-0.1}$	$0.6^{+0.3}_{-0.2}$	$1.6^{+0.8}_{-0.6}$	4^{+2}_{-1}
			Sanitised features				
Supervised DNN	6^{+3}_{-2}	$0.0035^{+0.0022}_{-0.0009}$	$0.006^{+0.003}_{-0.002}$	$0.009^{+0.004}_{-0.003}$	$0.014^{+0.010}_{-0.005}$	$0.07^{+0.04}_{-0.03}$	$0.15^{+0.09}_{-0.05}$
H_T	100^{+40}_{-30}	$0.14^{+0.07}_{-0.04}$	$0.16^{+0.08}_{-0.05}$	$0.16^{+0.08}_{-0.05}$	$0.4^{+0.2}_{-0.1}$	$1.0^{+0.5}_{-0.3}$	$1.8^{+0.9}_{-0.6}$
Deep SVDD	60^{+30}_{-20}	$0.25^{+0.13}_{-0.08}$	$0.16^{+0.08}_{-0.04}$	$0.12^{+0.05}_{-0.03}$	$0.5^{+0.2}_{-0.1}$	$1.0^{+0.5}_{-0.3}$	$2.0^{+0.8}_{-0.5}$
AE	160^{+60}_{-50}	$0.0099^{+0.0009}_{-0.0007}$	$0.0122^{+0.0006}_{-0.0009}$	$0.0152^{+0.0009}_{-0.0007}$	$0.0165^{+0.0007}_{-0.0011}$	$0.073^{+0.004}_{-0.004}$	$0.27^{+0.02}_{-0.02}$
HBOS	110^{+50}_{-30}	$0.19^{+0.11}_{-0.06}$	$0.21^{+0.12}_{-0.07}$	$0.23^{+0.14}_{-0.08}$	$0.4^{+0.2}_{-0.1}$	$1.1^{+0.7}_{-0.4}$	$2.7^{+1.7}_{-0.9}$
iForest	140^{+60}_{-40}	$0.3^{+0.2}_{-0.1}$	$0.4^{+0.2}_{-0.1}$	$0.4^{+0.2}_{-0.1}$	$0.8^{+0.4}_{-0.3}$	$2.2^{+1.2}_{-0.7}$	5^{+3}_{-2}

Finding New Physics without learning about it

Results 4: Can we search for New Physics?