Finding new physics without learning about it: anomaly detection as a tool for searches at colliders

Patrícia Ferreira

Introduction

- There is the concern that a new physics signal is missed because the method used for the search is not sensitive to it
- The use of anomaly detection methods is a way to increase sensitivity to new physics signals
- Anomaly detection methods are trained with Standard Model (SM) events and don’t depend on the physics we are trying to find

Shallow Methods

- **Histogram-based outlier detection (HBOS)**

creates histograms for each feature

for the test data, for each feature we see in what bin the value falls on

anomaly score calculated using log of the height of the bin and summing across all features

- **Isolation forest (iForest)**

recursively separates the data with a random limit

outliers are easier to isolate than normal events

Deep Methods

- **Deep autoencoder (AE)**

learns to compress and then decompress data

reconstruction error used as anomaly score

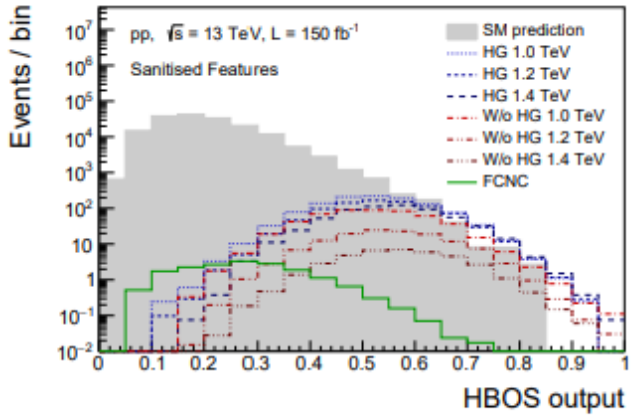
- **Deep Support vector data description (Deep SVDD)**

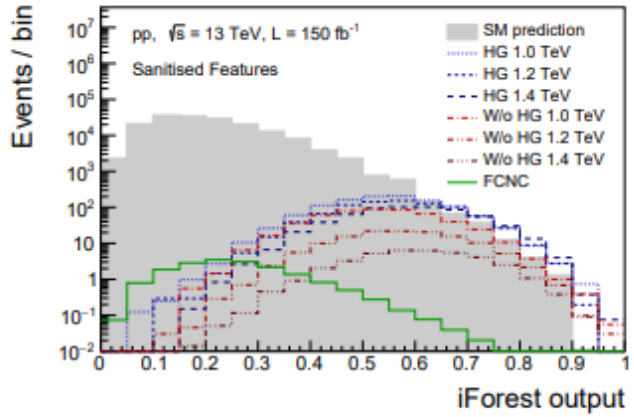
a DNN is initialized and the centre of the distribution is calculated

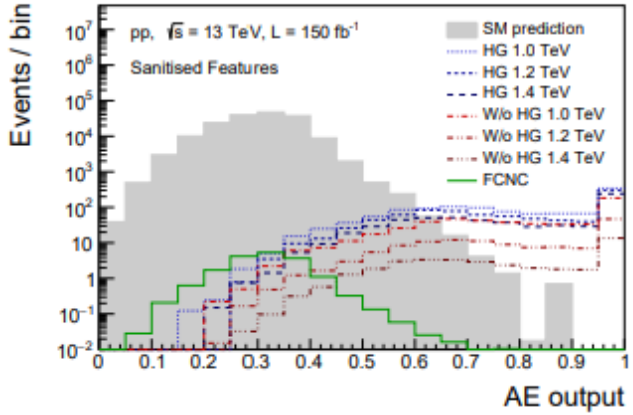
trained to minimize the distance of all points to the centre

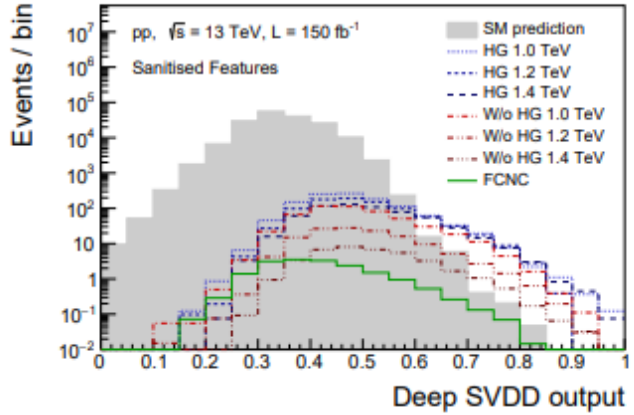
Anomaly score distribution

- FCNC signal distribution very close to the one followed by the SM processes
- Distributions for the signals shifted to the right, meaning that on average abnormal signals have higher anomaly scores than the SM events









Upper limits normalized to the supervised DNN

- Deep SVDD has similar sensitivity for all signals even for the FCNC
- Different AD algorithms are capturing outliers differently
- These AD algorithms are reasonably sensitive to new signals
- Degradation relative to the supervised DNN of around an order of magnitude for the worst cases, and no significant impact for the best ones