Search for heavy scalar resonances in the LHC run 2 dataset in $Z\gamma$ final states using machine learning techniques.
Overview

- We possibly have some signal in the 240-280 GeV region, but it’s too small and mixed with a huge background.
- We are setting up a search based on machine learning tools – This can split better the signal and background
- Conduct a weakly-supervised and full-supervised learning in conjunction with a Deep Neural Network Algorithm.

✔ Here is an illustration of the discovery of Higgs Boson, showing how in general HEP people are doing.
✔ We are trying to find a signal peak from such kind of a continuos background.
✔ A simillar problem we are dealing with!
Z + gamma

- We want to search for a heavy Higgs ($H$) decaying to $Z\gamma$ in association with $E_T^{\text{miss}}$.
- An illustration of one of the possible processes which is a heavy mass of particle $A$ decaying to $Z$ and $H$
Machine Learning (methodology)

- Full-supervised learning: requires distinct Signal and a Background samples (labelled data)

- Weakly-supervised learning: requires a Background sample and a sample from data that contains Background and an unknown mixture of Signal (unlabelled data), hidden in the real data.

Toolkit for Multivariate Data Analysis (TMVA) framework developed by CERN for high energy physics

MODEL CONFIGURATION: DNN_CPU

- One input layer with variation number of nodes
- 3 hidden layers with each 128 nodes
- One output layer
- Hyperbolic tangent activation function (TANH)
- BatchSize = 256
- Number of epochs = 80
Zγ mass spectrum

ATLAS Simulation Work in Progress

EVENTS / 2 GeV

M_{Zγ} [GeV]

Signal MW

M1: (Bkg side band)

M2: (Bkg MW + Sig MW)
Cuts requirement

- Signal + Background in Mass Window:
  - $m_{Zy} \leq 280 \&\& m_{Zy} \geq 240$ GeV

- Background side band:
  - $m_{Zy} \leq 240 \&\& m_{Zy} \geq 200 \| 
  - $m_{Zy} \leq 400 \&\& m_{Zy} \geq 280$ GeV

\[ E_{T, \text{sig}}^{\text{miss}} \geq 2.5 \text{ GeV} \]
1D distribution of input variables

M1: Signal MW + Background MW
M2: Background side band
1D distribution cont...

M1: Signal MW + Background MW
M2: Background side band

ATLAS Simulation
Work in Progress
1D distribution cont...

M1: Signal MW + Background MW
M2: Background side band
Overtraining check for DNN_CPU classifier

Weak-supervision

M1: Signal MW + Background MW

M2: Background side band

Training: 60502 events
Testing: 51535 events

Full-supervision

Training: 25767 events
Testing: 25767 events
ROC curve plot for DNN classifier

Weak-supervision

Full-supervision

ATLAS Simulation

Work in Progress

ROC-integ: 0.715

ROC-integ: 0.773
Conclusion

• A search for heavy Higgs in the Zy decay channel has been established. This was motivated by the statistically significant excesses in the multi-lepton final states compatible with physics at the EW scale.
• The search is based on machine learning techniques for separation of the signal from background - Toolkit for Multivariate Data Analysis with ROOT (TMVA).
• The combination of DNNs, advanced machine learning techniques and weak supervision methodology can help for finding new particles in High Energy Physics
• What we have tested in the Zy decay channel showed some improvements and a proof that indeed the search strategy is working.
• This methodology can be implemented in the LHC searches and analyses towards discovery of new physics.
THANK YOU
Back-up slides (1D distribution of input variables- Full supervision)