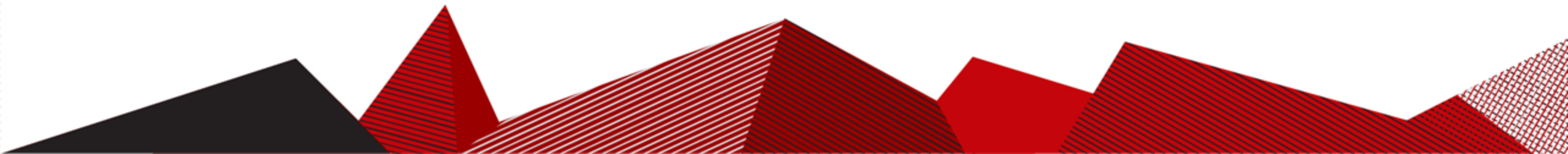
Feasibility Study of Measuring the Higgs Self-coupling Using the Muon Collider

Kenny Jia

Sep 2nd, 2021

B-tagging for $k_T/anti\ k_T$



Adding JetFlavorAssociation module for $anti - k_T$

cards/MuonCollider/MuonColliderDet_ JetFlavorAssociation.tcl

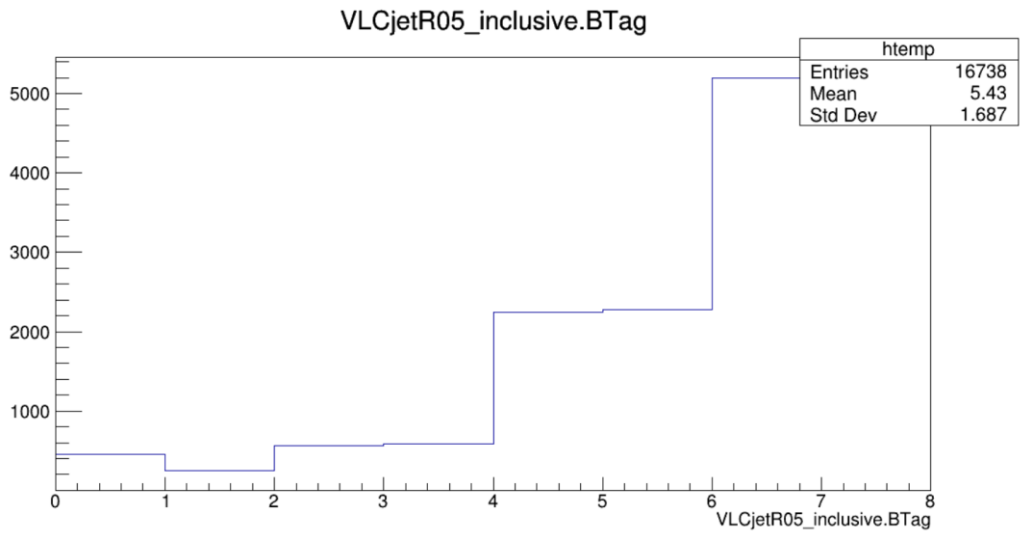
```
############################
# AKT JetFlavorAssociation
############################
module JetFlavorAssociation AKT_JetFlavorAssociation_R05_inclusive {

    set PartonInputArray Delphes/partons
    set ParticleInputArray Delphes/allParticles
    set ParticleLHEFInputArray Delphes/allParticlesLHEF
    set JetInputArray FastJetFinderAKt/AKTjets

    set DeltaR 0.5
    set PartonPTMin 1.0
    set PartonEtaMax 2.5

}
```

Add new modules to Execution process in the main card

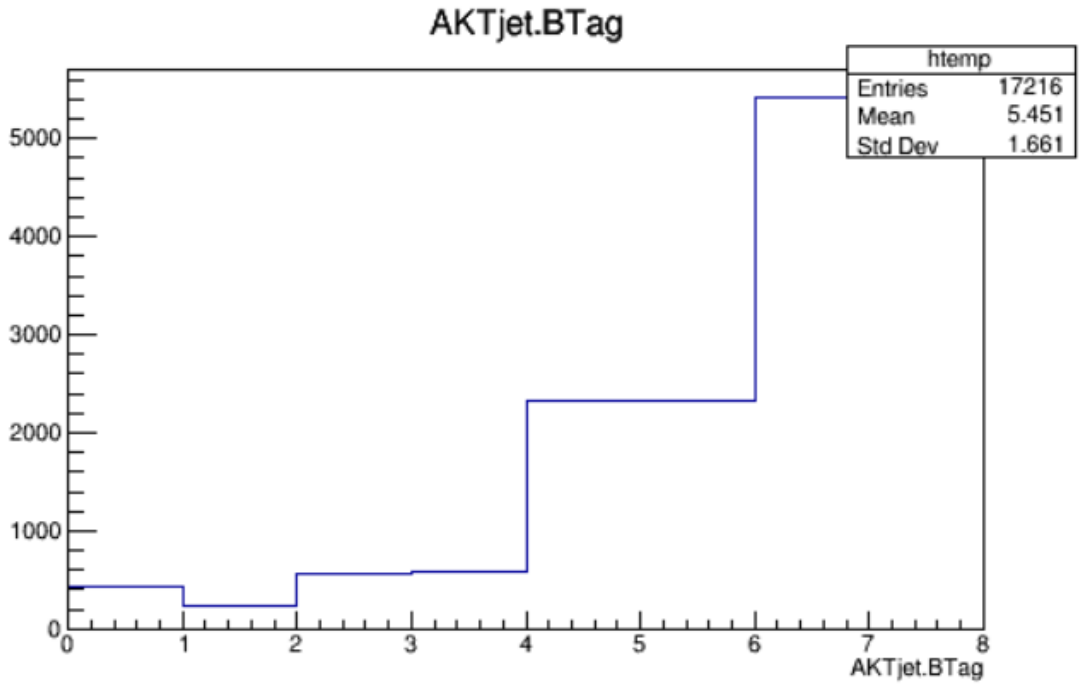
```
JetFlavorAssociation_R02_inclusive
JetFlavorAssociation_R05_inclusive
JetFlavorAssociation_R07_inclusive
JetFlavorAssociation_R10_inclusive
JetFlavorAssociation_R12_inclusive
JetFlavorAssociation_R15_inclusive

AKT_JetFlavorAssociation_R05_inclusive

AKT_BTagging_WP50_R05_inclusive
AKT_BTagging_WP70_R05_inclusive
AKT_BTagging_WP90_R05_inclusive
```

Comparison with original algo for 2v4b(10k)

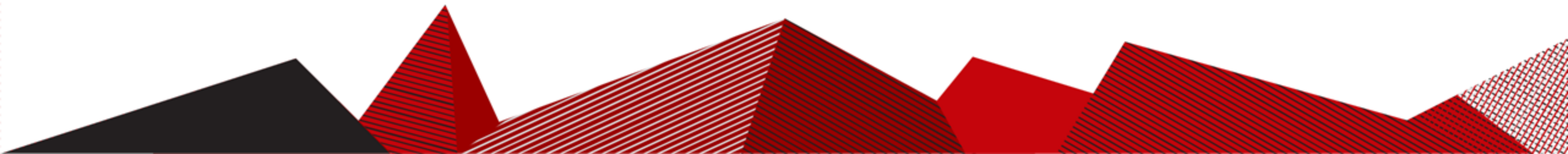
VLC R05_inclusive BTag



Anti kT R05 BTag



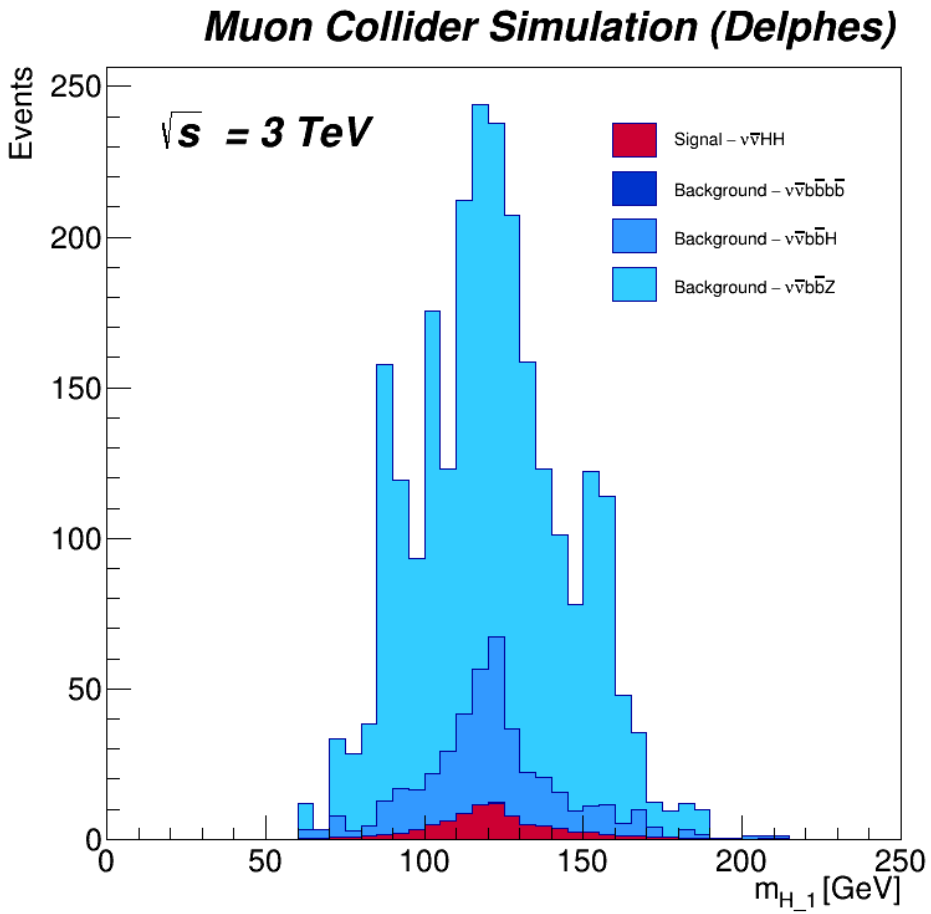
MVA for discriminating sig and bkg

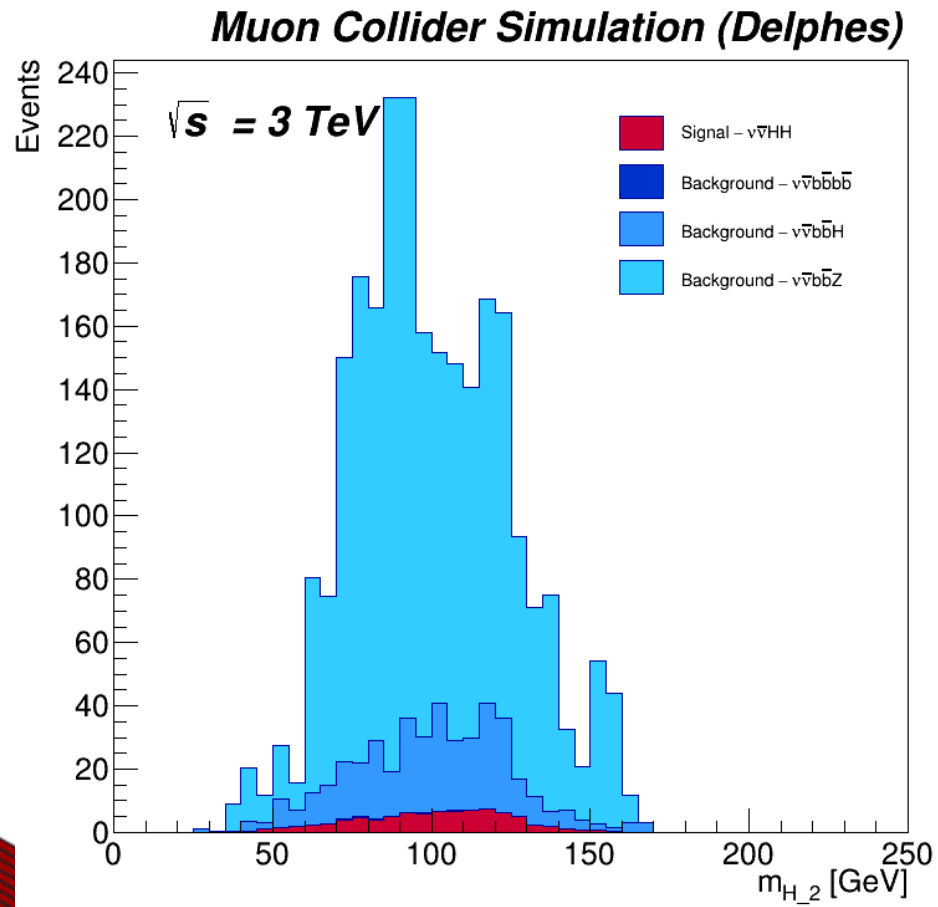


- Signal: $\mu^- + \mu^+ \rightarrow \nu_\mu + \bar{\nu}_\mu + H + H$ $(0.0008201 \pm 2.2\text{e} - 6\ \text{pb})$
- Background:
  - $\mu^- + \mu^+ \rightarrow \nu_\mu + \bar{\nu}_\mu + b + \bar{b} + Z$ $(0.03168 \pm 0.00027\ \text{pb})$
  - $\mu^- + \mu^+ \rightarrow \nu_\mu + \bar{\nu}_\mu + b + \bar{b} + H$ $(0.003743 \pm 2.7\text{e} - 5\ \text{pb})$
  - $\mu^- + \mu^+ \rightarrow \nu_\mu + \bar{\nu}_\mu + b + \bar{b} + b + \bar{b}$ $(0.0009227 \pm 7.02\text{e} - 6\ \text{pb})$

10k for each

Normalized by cross section





MVA to discriminate vvHH with dominant bkg vvbbZ

- With TMVA, a built-in package in ROOT
  - Support different ML approach:
    - Boosting Decision Tree
    - Multilayer perceptron (a type of feed-forward Artificial Neural Network)
    - Deep Neural Network
  - External Machine learning framework supported:
    - XGBoost (I used in CMSDAS but not through TMVA), keras
- Transfer to python, use other external framework like TensorFLow or PyTorch
  - Need to switch Root data to panda or numpy
  - Need installation

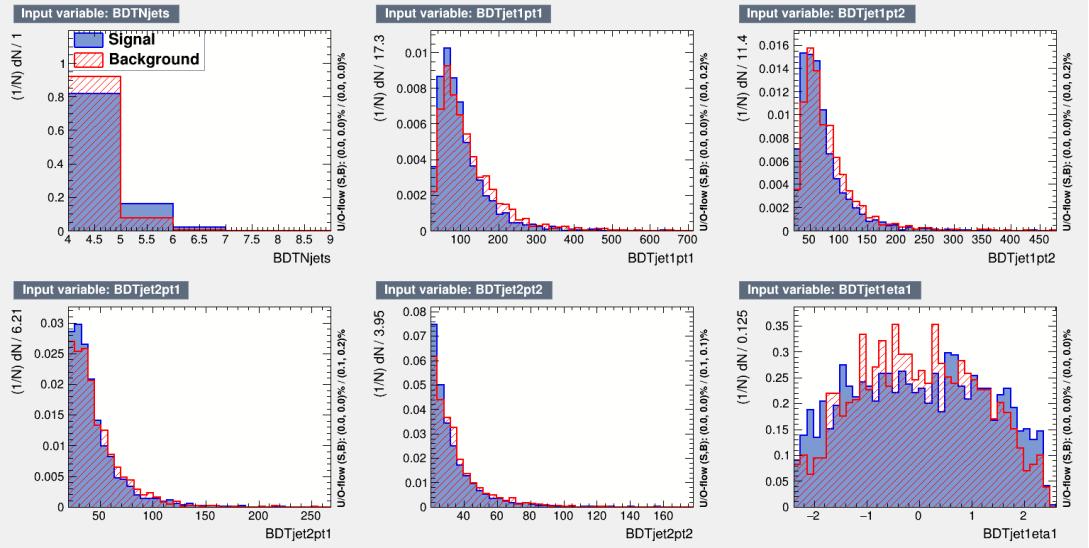
Trained variables

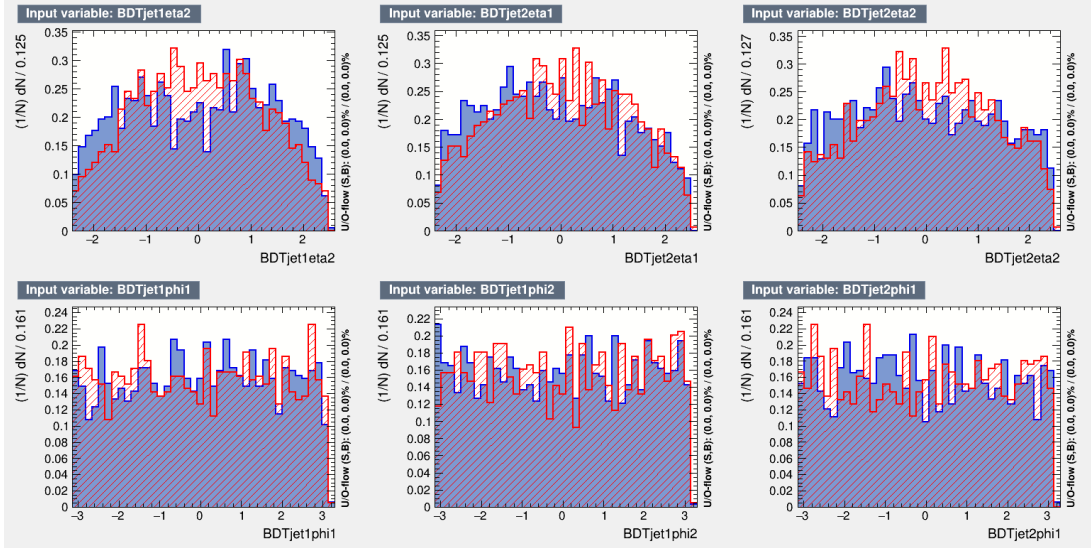
- Number of jets in the event
- p_T for all four jets, higgs, and di-Higgs system
- η and ϕ for all four jets, higgs, and di-Higgs system
- Invariant mass for Higgs and di-Higgs system
- B-Tagging for all four jets, number of b-tagged jets in the event
- MET, η and ϕ of MET
- deltaTheta between di-higgs and MET
- deltaPhi, deltaEta for higgs

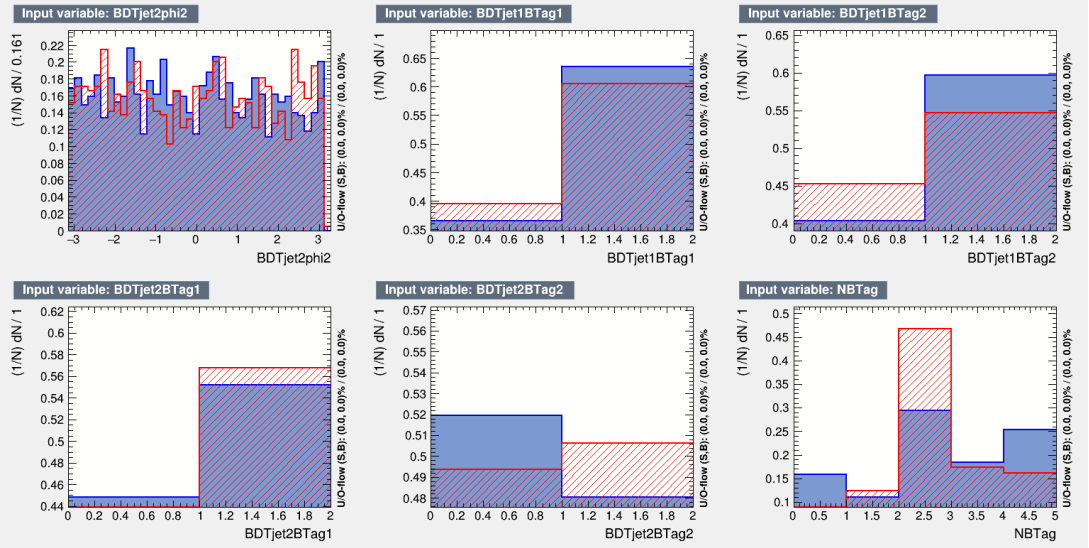
Trained variables

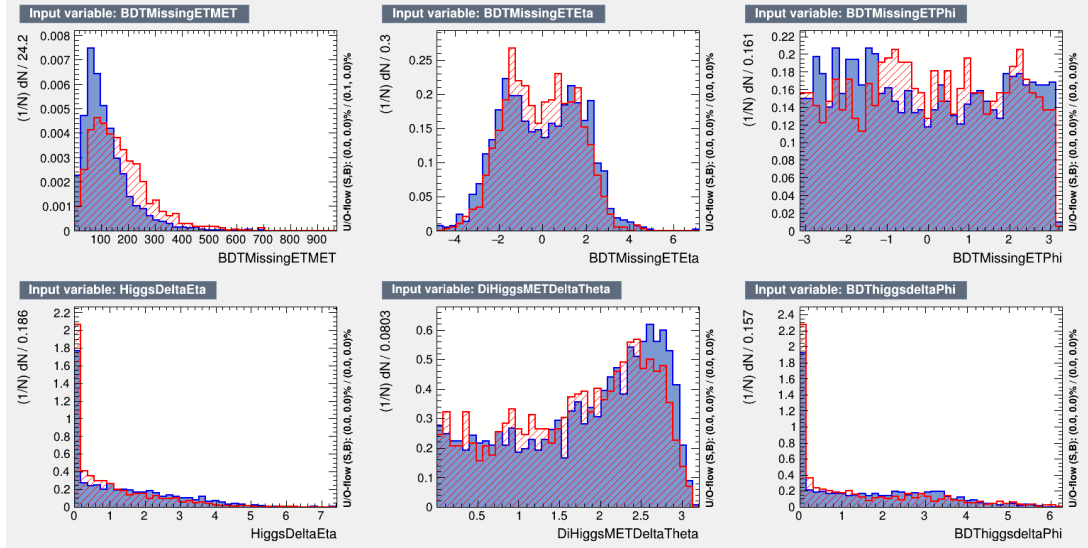
“With ”Useful” set performing worse than ”Useless” set, this surprising result seems to imply that the number of variables used in BDT training is more important than the choices of which variables to use.”

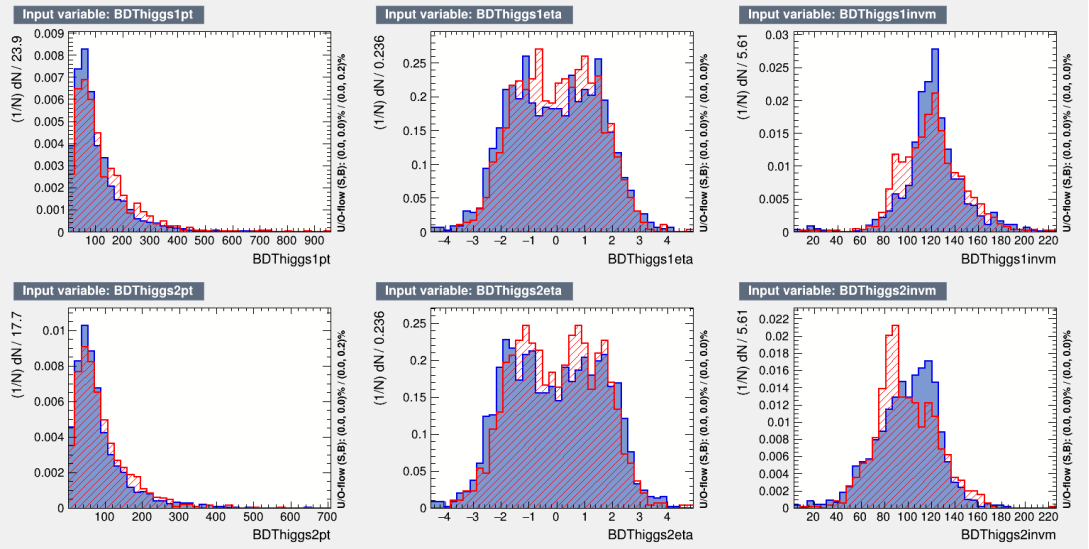
From “Optimizing Variable Selection for Machine Learning Analysis in ATLAS ttH Search”

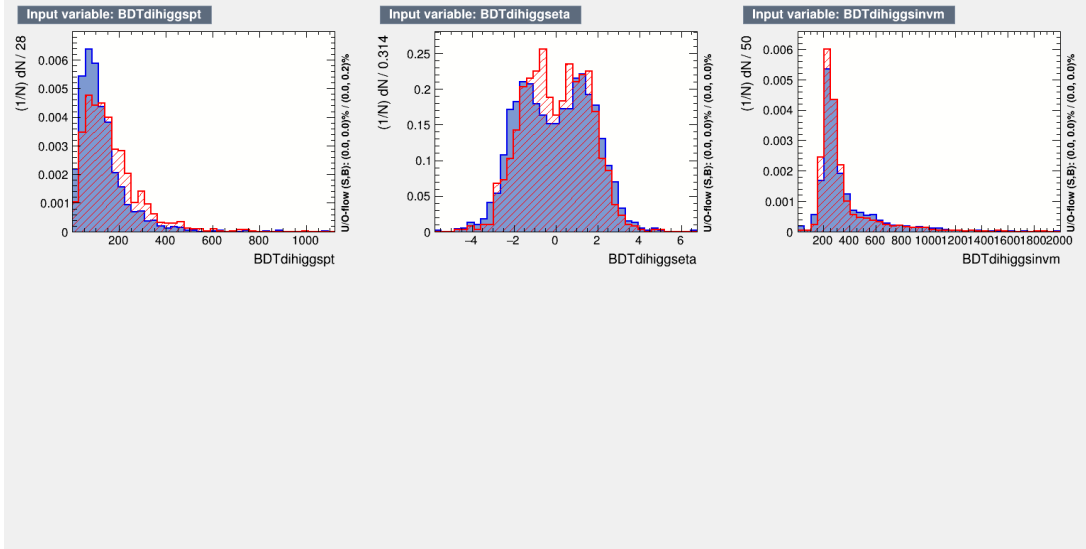

Input variable: BDTNjets
Signal
Background
(1/N) dN / 1
BDTNjets
Input variable: BDTjet1pt1
(1/N) dN / 17.3
BDTjet1pt1
Input variable: BDTjet1pt2
(1/N) dN / 11.4
BDTjet1pt2
Input variable: BDTjet2pt1
(1/N) dN / 6.21
BDTjet2pt1
Input variable: BDTjet2pt2
(1/N) dN / 3.95
BDTjet2pt2
Input variable: BDTjet1eta1
(1/N) dN / 0.125
BDTjet1eta1
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%


Input variable: BDTjet1eta2
(1/N) dN / 0.125
BDTjet1eta2
Input variable: BDTjet2eta1
(1/N) dN / 0.125
BDTjet2eta1
Input variable: BDTjet2eta2
(1/N) dN / 0.127
BDTjet2eta2
Input variable: BDTjet1phi1
(1/N) dN / 0.161
BDTjet1phi1
Input variable: BDTjet1phi2
(1/N) dN / 0.161
BDTjet1phi2
Input variable: BDTjet2phi1
(1/N) dN / 0.161
BDTjet2phi1
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%


Input variable: BDTjet2phi2
(1/N) dN / 0.161
BDTjet2phi2
Input variable: BDTjet1BTag1
(1/N) dN / 1
BDTjet1BTag1
Input variable: BDTjet1BTag2
(1/N) dN / 1
BDTjet1BTag2
Input variable: BDTjet2BTag1
(1/N) dN / 1
BDTjet2BTag1
Input variable: BDTjet2BTag2
(1/N) dN / 1
BDTjet2BTag2
Input variable: NBTag
(1/N) dN / 1
NBTag
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%


Input variable: BDTMissingETMET
(1/N) dN / 24.2
BDTMissingETMET
Input variable: BDTMissingETEta
(1/N) dN / 0.3
BDTMissingETEta
Input variable: BDTMissingETPhi
(1/N) dN / 0.161
BDTMissingETPhi
Input variable: HiggsDeltaEta
(1/N) dN / 0.186
HiggsDeltaEta
Input variable: DiHiggsMETDeltaTheta
(1/N) dN / 0.0803
DiHiggsMETDeltaTheta
Input variable: BDThiggsdeltaPhi
(1/N) dN / 0.157
BDThiggsdeltaPhi
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

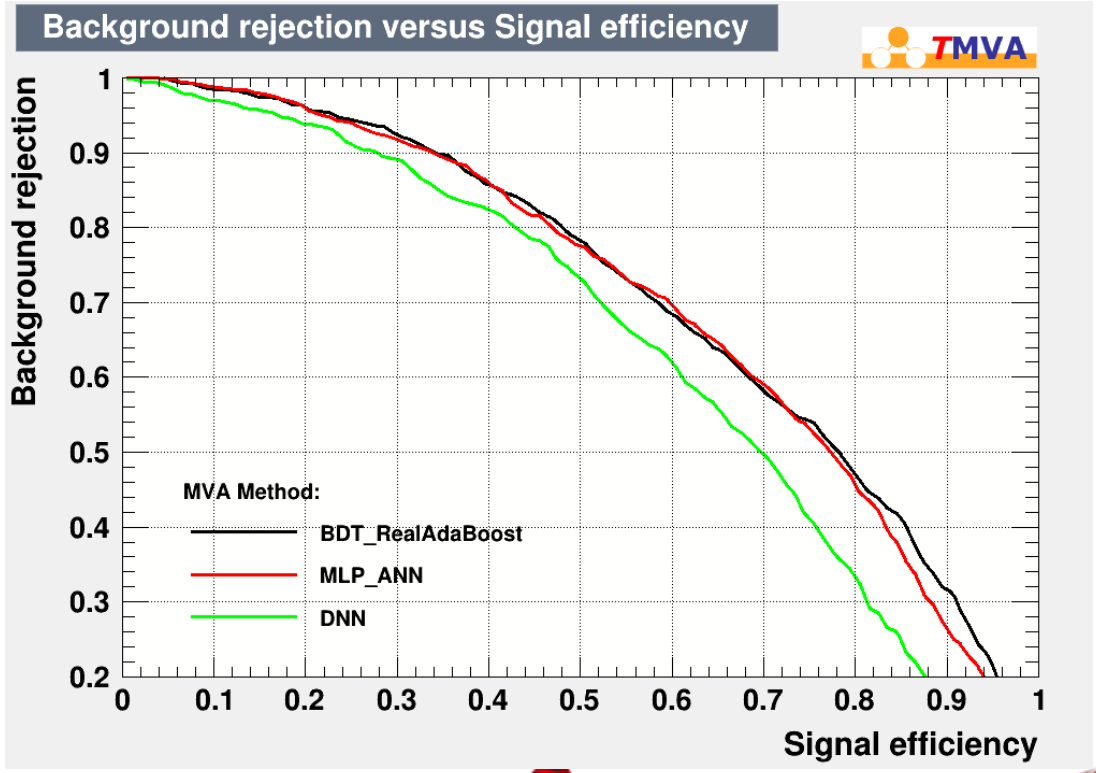

Input variable: BDThiggs1pt
(1/N) dN / 23.9
BDThiggs1pt
Input variable: BDThiggs1eta
(1/N) dN / 0.236
BDThiggs1eta
Input variable: BDThiggs1invm
(1/N) dN / 5.61
BDThiggs1invm
Input variable: BDThiggs2pt
(1/N) dN / 17.7
BDThiggs2pt
Input variable: BDThiggs2eta
(1/N) dN / 0.236
BDThiggs2eta
Input variable: BDThiggs2invm
(1/N) dN / 5.61
BDThiggs2invm
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.2)%
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

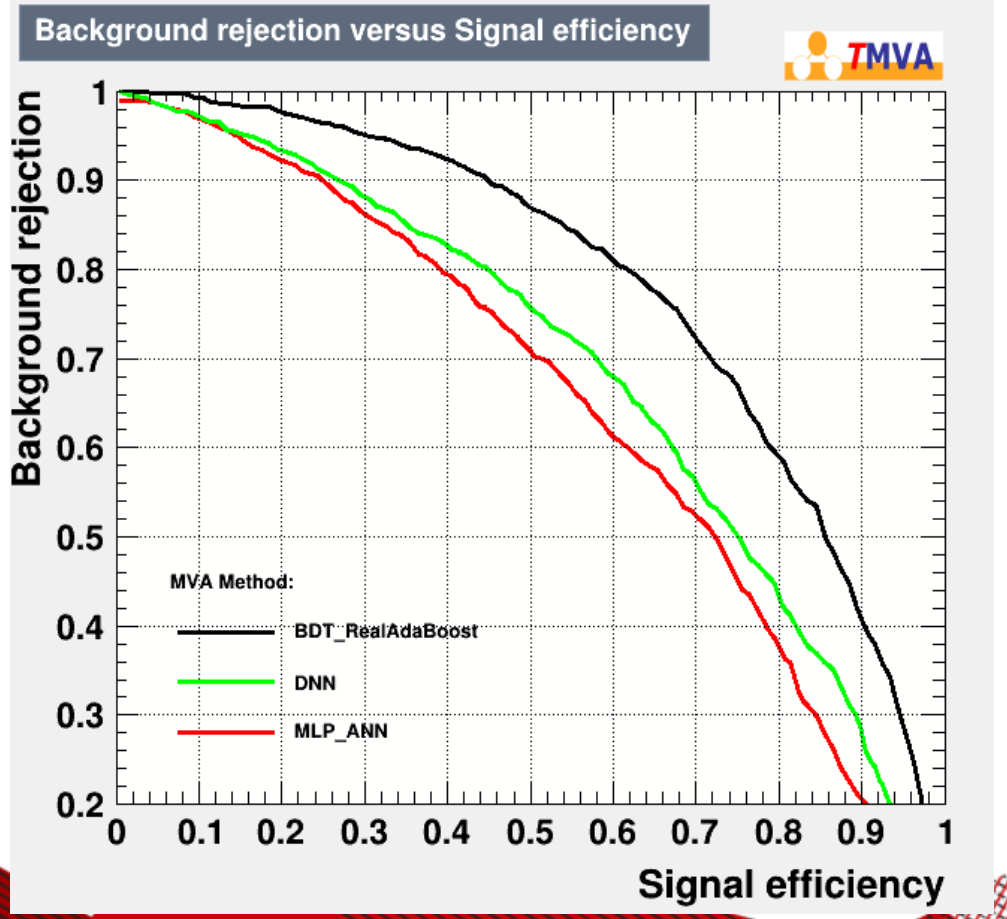

Input variable: BDTdihiggspt
(1/N) dN / 28
BDTdihiggspt
Input variable: BDTdihiggseta
(1/N) dN / 0.314
BDTdihiggseta
Input variable: BDTdihiggsinvm
(1/N) dN / 50
BDTdihiggsinvm
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.2)%
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

Boosting Dicision Trees

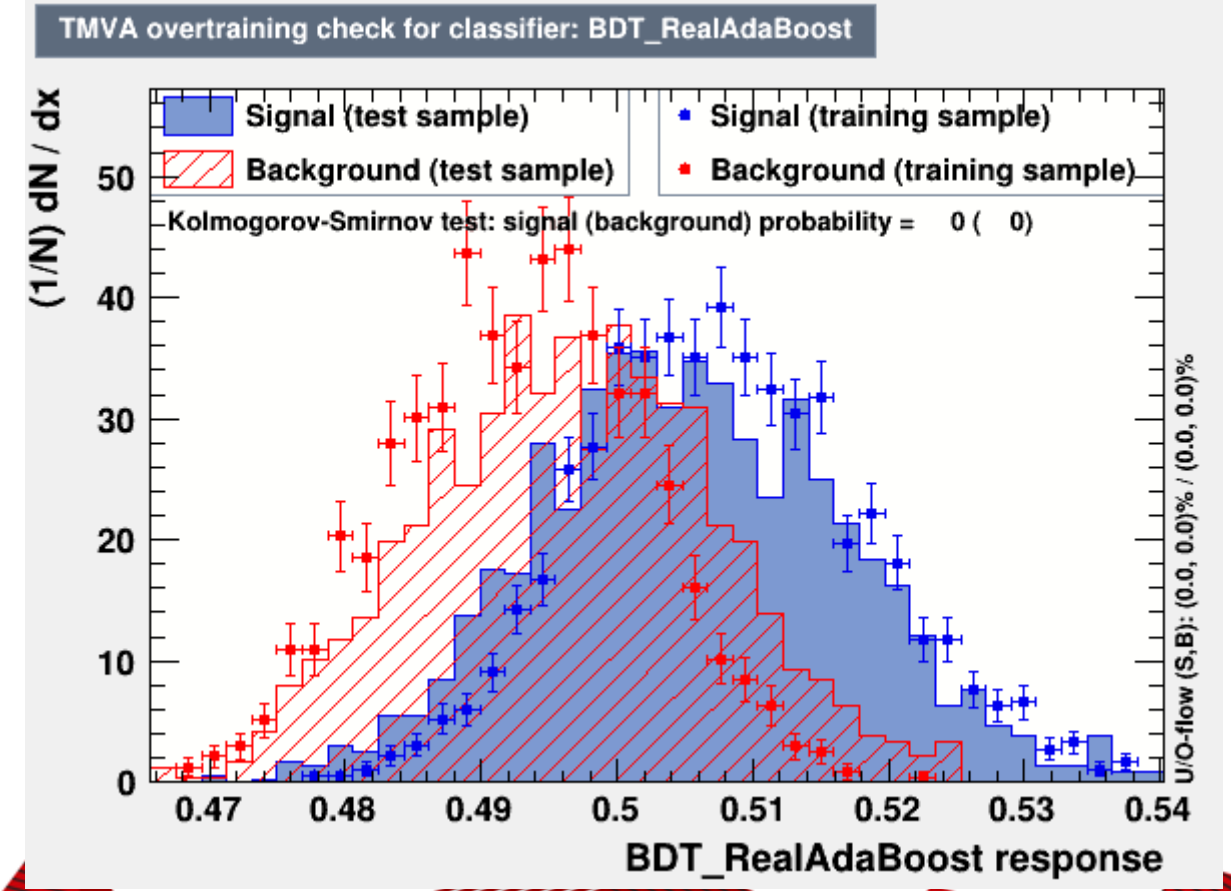
- 400 trees, MaxDepth = 3
- BoostTypr = RealAdaBoost,
- Separation Type: Gini Index
- No pruning since we don’t worry about overtraining for our case

AUC improved from 0.713 to 0.784

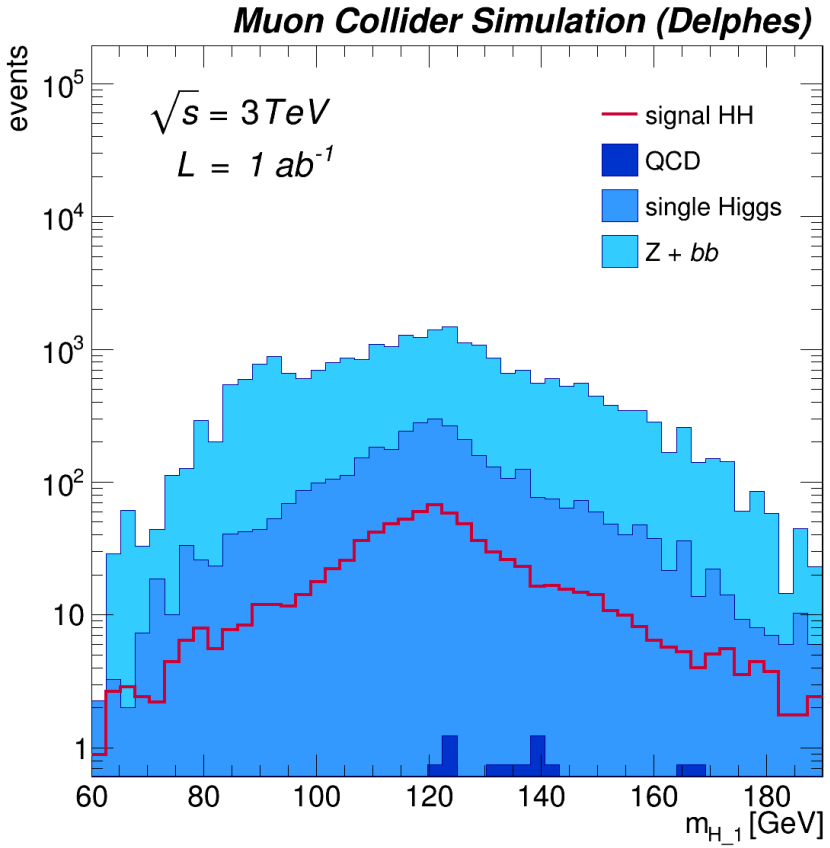


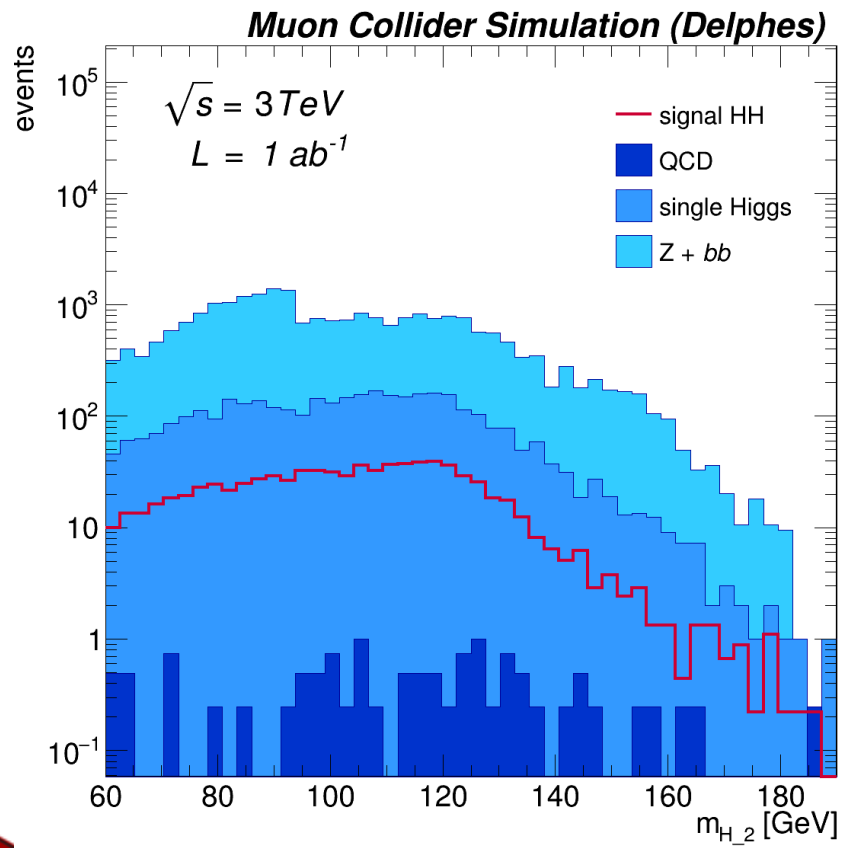


Response for BDT_RealAdaBoost



Without any BDT cut





BDT Application