Study with new proposed features for various AE

Rishabh Uniyal
Catholic U, Washington DC
August 26, 2019

PD Features mapping

new features (electron in EGamma)
Features removed

ZeroBias	JetHT	EGamma	SingleMuon
qpVtxChi2 qpVtxNtr qpVtxX qpVtxY qpVtxZ qPUEvt qlumiEvt qgTkPt qgTkEta qgTkPhi qgTkN qgTkChi2 qgTkNHits qgTkNLay 7*13 = 91 features	qpVtxChi2 qpVtxNtr qpVtxX qpVtxY qpVtxZ qPUEvt qlumiEvt PFJetN PFJetPt PFJetPhi PFJetEta PFMetPt PFMetPhi CalJetN CalJetPt CalJetEta CalJetPhi CalJetEn CalMETPt CalMETPhi CCEn_ CCEta_ CCPhi_ SCEn_ SCEta_ SCPhi_ 7*24 = 168	qpVtxChi2 qpVtxNtr qpVtxX_ qpVtxY_ qpVtxZ_ qPUEvt_ qlumiEvt **qGsfPt** **qGsfEta** **qGsfPhi** **qGsfN** qPhoN gedPhoPt gedPhoEta gedPhoPhi gedPhoEn gedPhoe1x5 gedPhoe3x3 **SigmaIEta_; SigmaIPhi_; r9_; HadOEm_; drSumPt_; drSumEt_; eSCOP_; ecEn_;** 7*30 = 210	qpVtxChi2_ qpVtxNtr_ qpVtxX_ qpVtxY_ qpVtxZ_ qPUEvt_ qlumiEvt_ qglobTkN qglobTkPt qglobTkEta qglobTkPhi _ qglobTkChi2 qglobTkNHits **qMuNCh** qMuN qMuPt qMuEta qMuPhi qMuEn qMuChi2 7*18 = 126

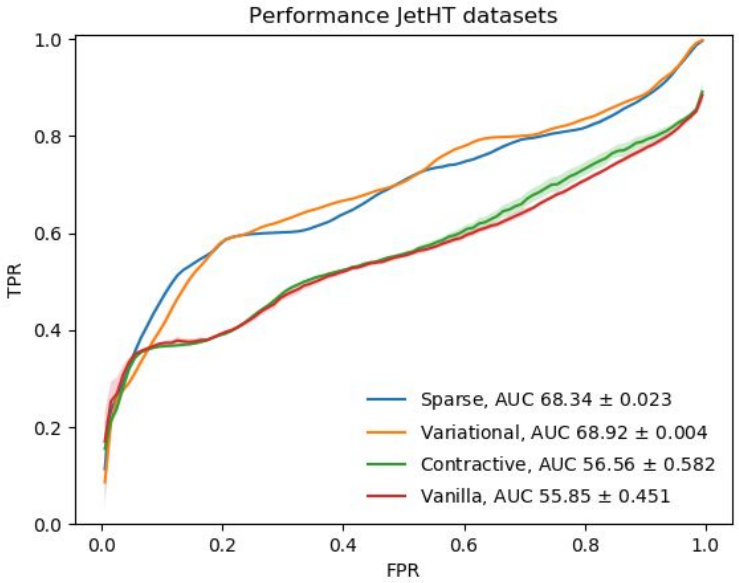
2

JetHT

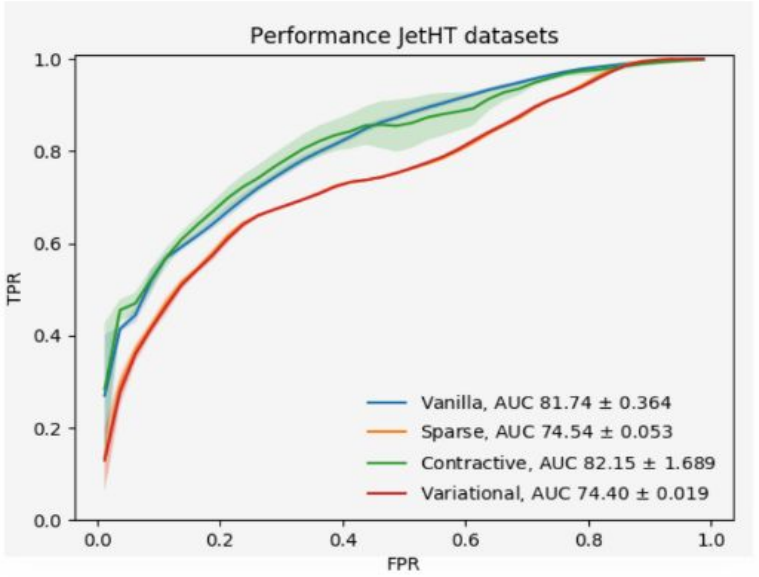
of Epochs - 1200
BS = 2e15
Dataset 2018

Performance goes down with new features

With changed features



Previously used features by Jab



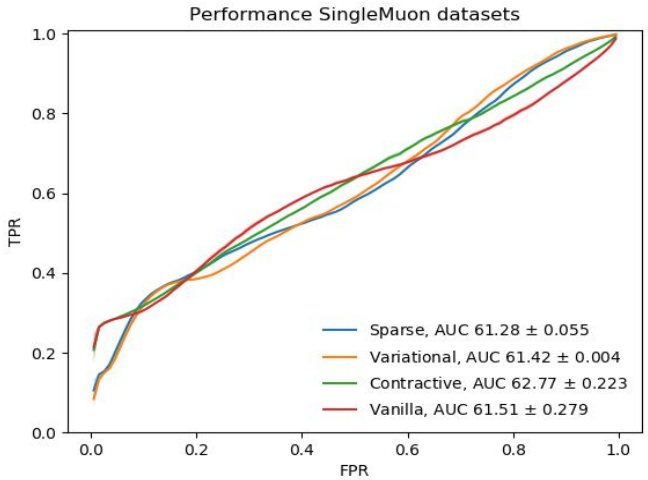
https://github.com/calzonelover/CMS_DC_ANOMALY/blob/master/report/reco/new_data/reports/22july2019.pdf

SingleMuon

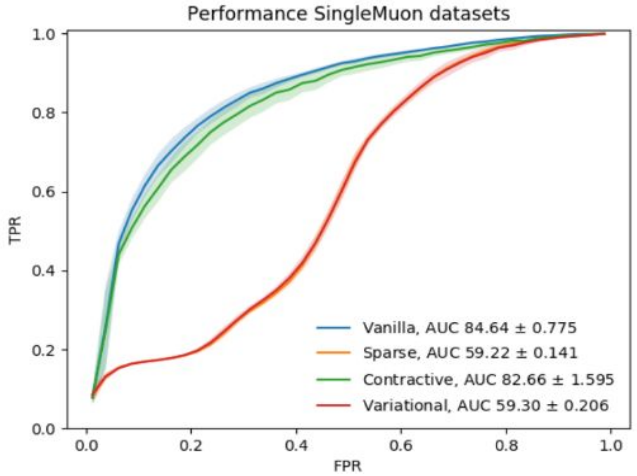
of Epochs - 1200
BS = 2e15
Dataset 2018

Performance goes down with new features

With changed features



Previously used features by Jab



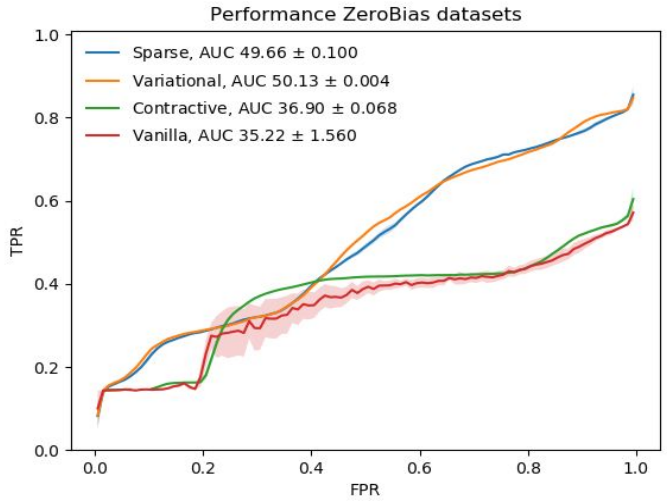
https://github.com/calzonelover/CMS_DC_ANOMALY/blob/master/report/reco/new_data/reports/22july2019.pdf

ZeroBias

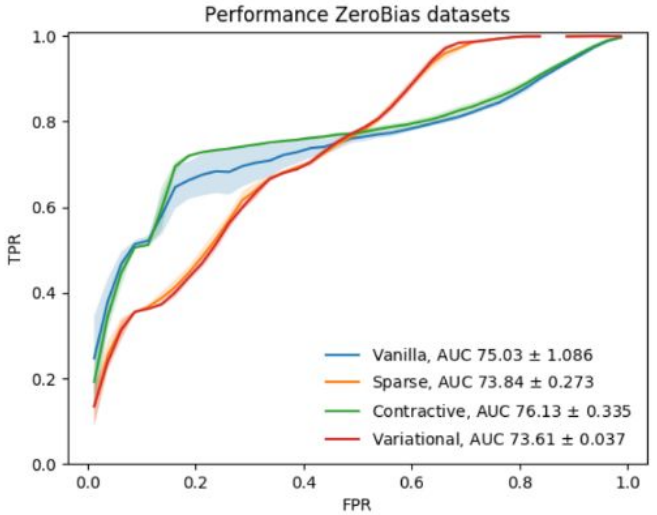
of Epochs - 1200
BS = 2e15
Dataset 2018

Performance goes down with new features

With changed features



Previously used features by Jab



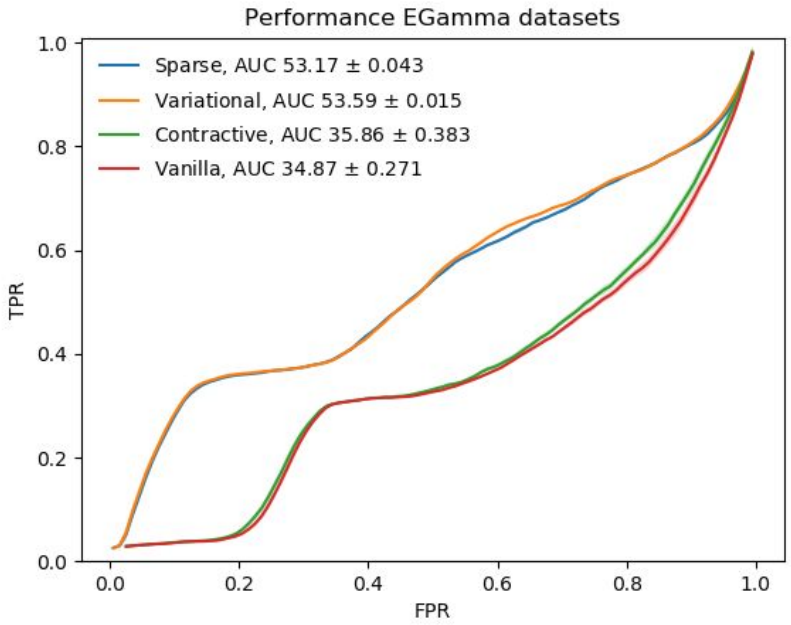
https://github.com/calzonelover/CMS_DC_ANOMALY/blob/master/report/reco/new_data/reports/22july2019.pdf

EGamma

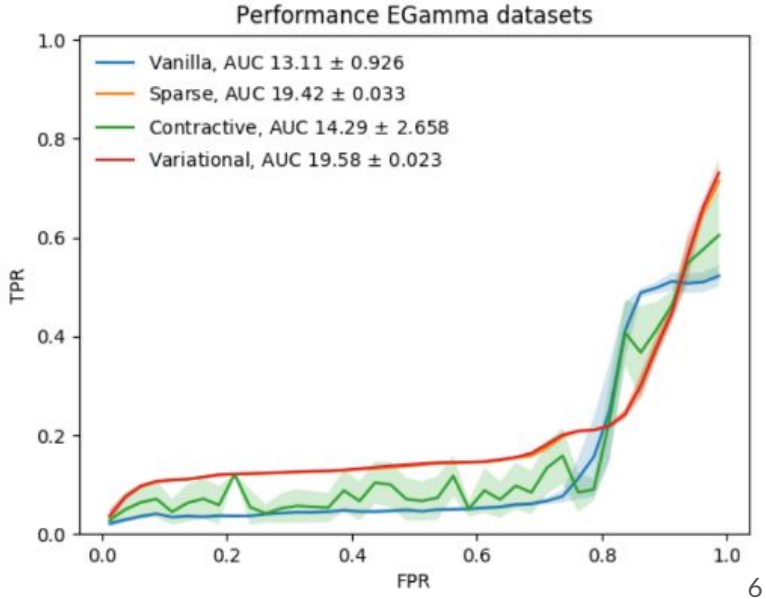
of Epochs - 1200
BS = 2e15
Dataset 2018

Performance improved for EGamma

With changed features

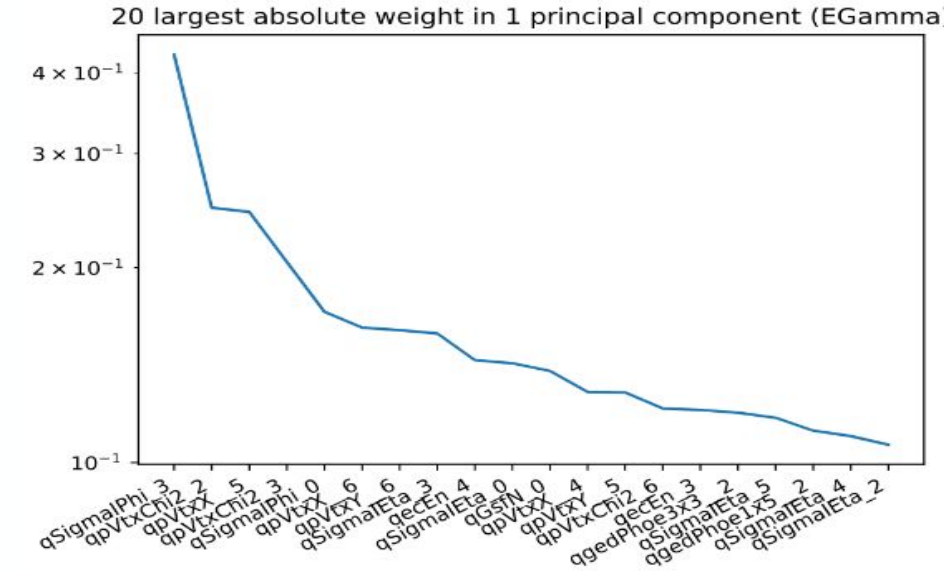


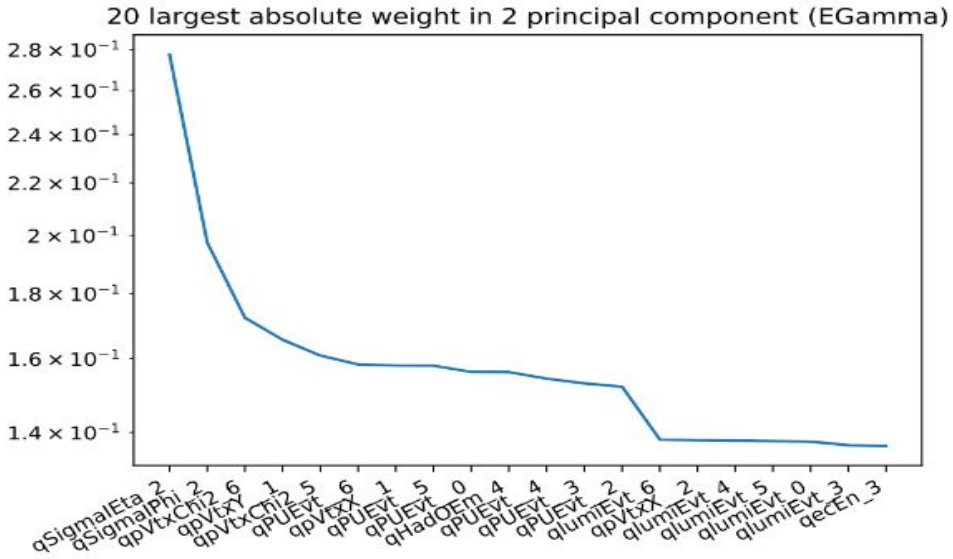
Previously used features Jab



https://github.com/calzonelover/CMS_DC_ANOMALY/blob/master/report/reco/new_data/reports/22july2019.pdf

Why AE performs better with new features for EGamma? (PCA analysis by Jab)





qpVtx and qpVtxY are the dominant features so, removing them should decrease the performance of the AE

Dominated features
- **qpVtxX and qpVtxY**
- **qSigmalEta**

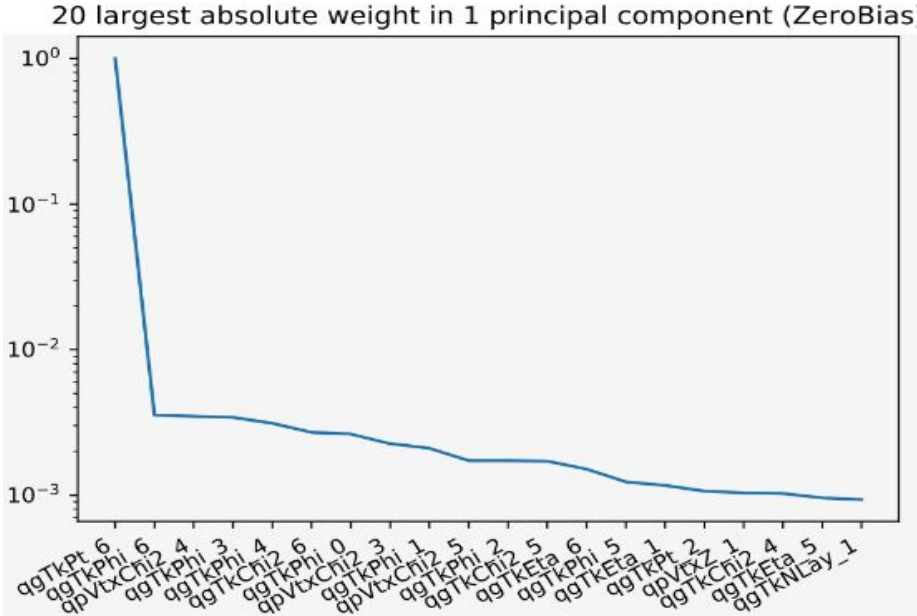
Overlapping feature
- **qSigmalPhi**
- **qpVtxChi2**

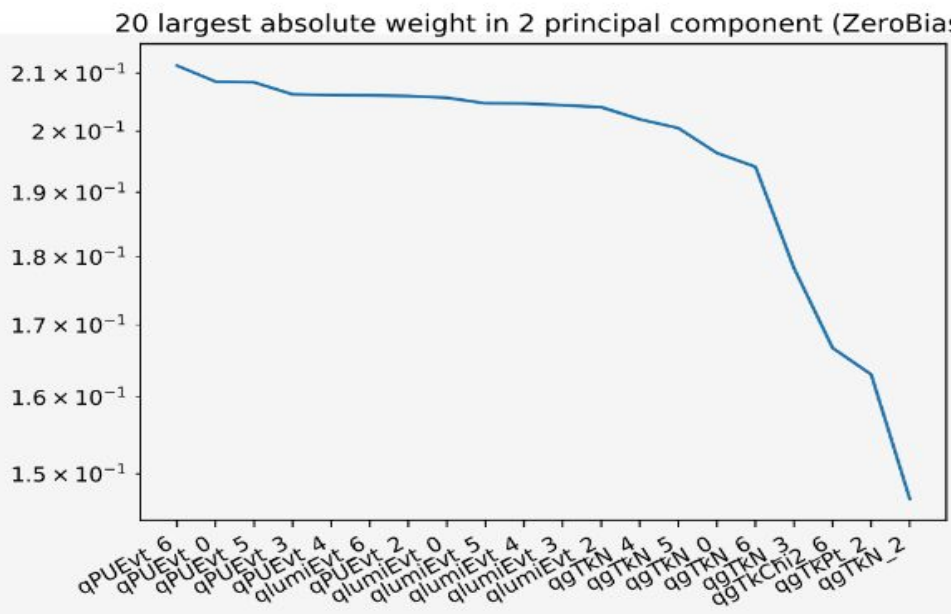
Dominated features
- **qPUEvt**
- **qlumiEvt**

Explained variance ratio ~ [0.31 0.25]

https://github.com/calzonelover/CMS_DC_ANOMALY/blob/master/report/reco/new_data/reports/12Aug2019.pdf

ZeroBias



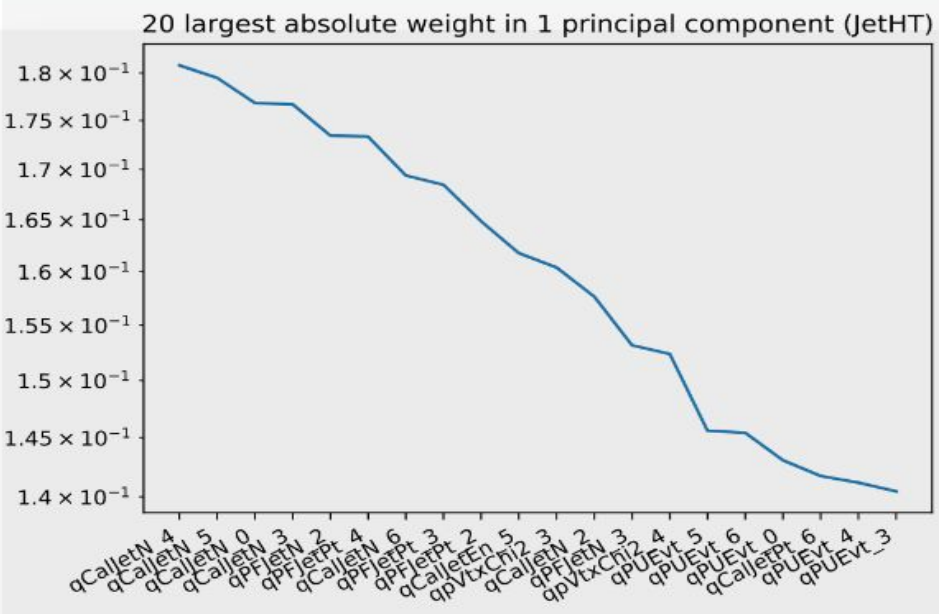


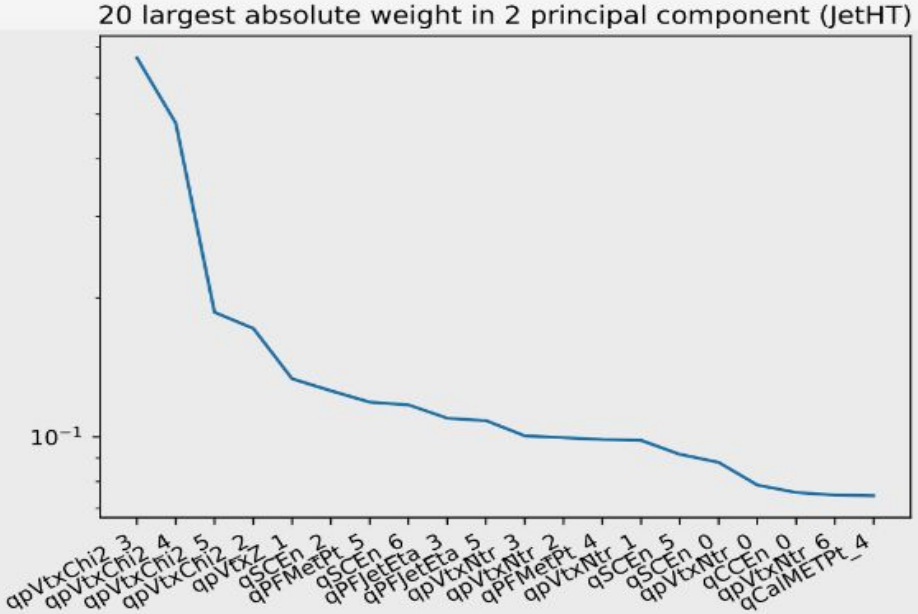
Dominated features
- **qgTkPt**
- **qgTkPhi**

Dominated features
- **qPUEvt**
- **qlumiEvt**
- **qgTkN**

https://github.com/calzonelover/CMS_DC_ANOMALY/blob/master/report/reco/new_data/reports/12Aug2019.pdf

JetHT

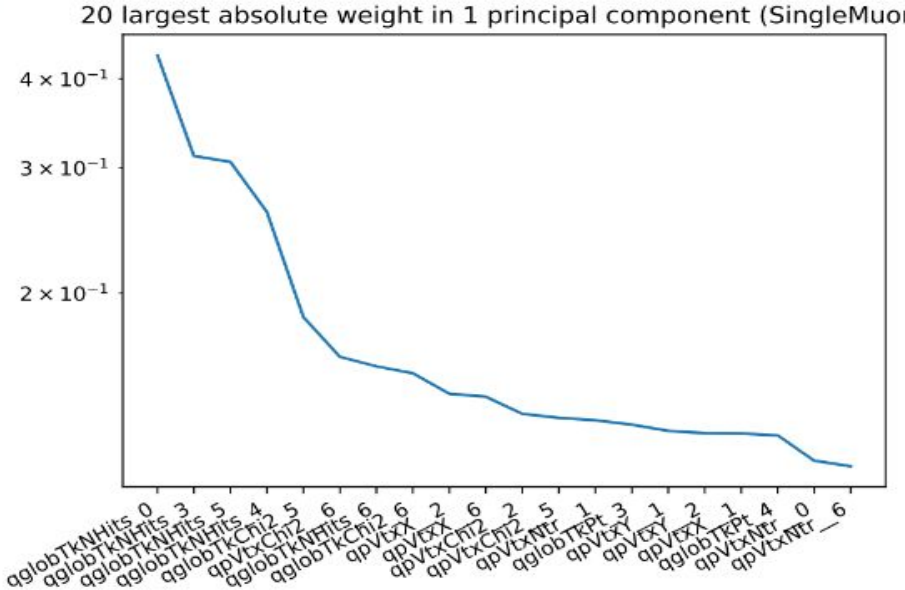


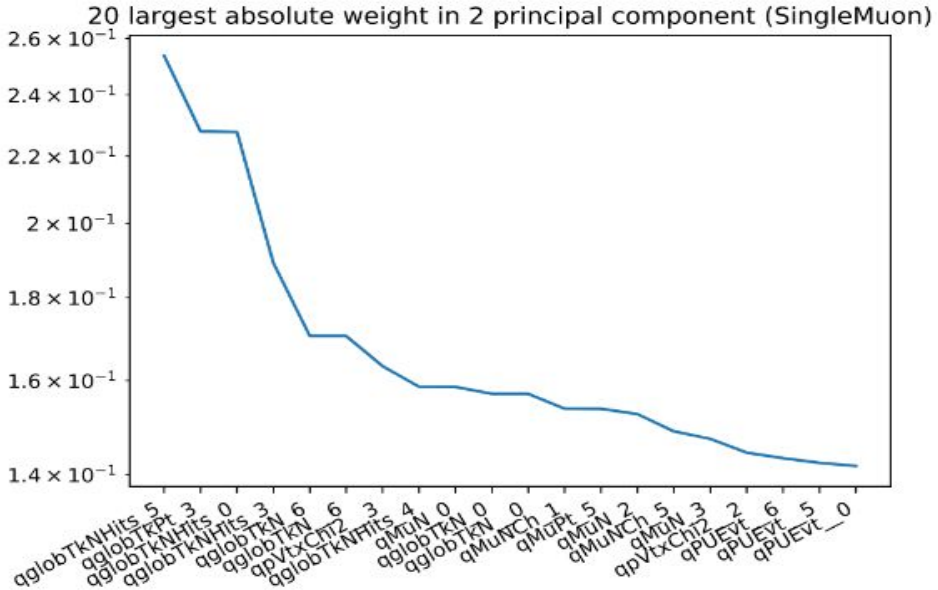


Dominated features
- **qCalJetN**
- **qCalJetPt**
- **qPUEvt**

Dominated features
- **qpVtxChi2**
- **qPFMetPt and qPFJetEta**

SingleMuon





Dominated features
- **qglobTkChi2**
- **qpVtxX and qpVtxY**

Overlapping feature
- **qglobTkNHits**

Dominated features
- **qPUEvt**
- **qMuN and qMuNCh**

Next steps??