Soft Muon Finder: Existing MVA and ID Overview

Margaret Lazarovits

April 15, 2020

1 Soft Muon ID

Developed by BHP group. See Table 1 for requirements going into the ID.

Input Variable	CMSSW Class
isGoodMuon	pat muon
recoMu.innerTrack()→hitPattern(). trackerLayersWithMeasurement() > 5	innerTrack.hitPattern
recoMu.innerTrack()→hitPattern(). pixelLayersWithMeasurement() > 0	innerTrack.hitPattern
recoMu.innerTrack()→quality(reco::TrackBase::highPurity)	TrackBase
fabs(recoMu.innerTrack()→dxy(vertex→position())) < 0.3 && fabs(recoMu.innerTrack()→dz(vertex→position())) < 20.	TrackBase

Table 1: Soft Muon ID requirements.

2 Soft Muon MVA

This MVA is a GBR (Gradient Boosted Regression) Forest built from TMVA-trained trees. Table 2 shows the preselections made before the GBRForest. Table 3 shows the variables that are fed into the GBRForest.

Preselection Requirement	CMSSW Class
global track is nonNull	muon.TrackRef
inner track is nonNull	muon.TrackRef
outer track is nonNull	muon.TrackRef

Table 2: Soft Muon MVA preselection table where muon.TrackRef is a reference to a collection of tracks in the muon class

Input Variable	Description	CMSSW Class
pt	muon pt	pat muon
eta	muon eta	pat muon
chi2LocalMomentum	chi2 value for the STA-TK matching of local momentum	muonQuality
chi2LocalPosition	chi2 value for the STA-TK matching of local position	muonQuality
glbTrackProbability	the tail probability (-ln(P)) of the global fit	muonQuality
trkRelChi2	chi2 value for the inner track stub with respect to the global track	muonQuality
trkKink	value of the kink algorithm applied to the inner track stub	muonQuality
glbKink	log(2 + glbKink); value of the kink algorithm applied to the global track	muonQuality
segmentCompatibility	segment compatibility for a track with matched muon info	pat muon
timeAtIpInOutErr	time of arrival at the IP for the Beta=1 hypothesis where particle is moving from outside in	muonTime
VMuonHitComb	DThits/2 + RPChits per muon station (up to 4)	the stored value is an int, but you need to count global hits
Inner Track: validFraction	fraction of valid hits on the track	inner track
Inner Track: normalizedChi2	χ^2 of track fit divided by n.d.o.f. (or $\chi^2 \times (1 \times 10^6)$ if n.d.o.f. is 0)	inner track
Inner Track: layers with measurement	number of tracker barrel layers with measurement	inner track.HitPattern
Outer Track: normalizedChi2	χ^2 of track fit divided by n.d.o.f. (or $\chi^2 \times (1 \times 10^6)$ if n.d.o.f. is 0)	outer track
Inner & Outer Tracks: QProd	product of track electric charge	inner and outer tracks

Table 3: Soft Muon MVA input variables

3 Lepton (Muon) MVA

Trained on inclusive ttZ LO Madgraph for signal (prompt leptons) and semi-leptonic ttbar for background (nonprompt leptons). Table 4 shows preselection requirements for the lepton and the associated jet. This MVA also uses a Gradient-Boosted Regression Forest similar to the Soft Muon MVA. See Twiki for more information on variables. The preselection for jets is done in MuonMvaEstimator.cc (but not for leptons - this is done in CMSSW muon preselection).

Preselection Variable Requirement	CMSSW Class
pt $>$ 5 GeV	pat muon
loose PF ID (isLooseMuon)	pat muon
mini-Isolation $<$ 0.4 (miniIsoLoose)	pat muon
SIP3D $<$ 8	pat muon
abs(dxy) $<$ 0.05 && abs(dz) $<$ 0.1	pat muon
track pt $>$ 1 GeV	PFCandidate
track charge() != 0	PFCandidate
deltaR w.r.t. jet eta,phi $\leq$ 0.4	I don't see this requirement either - deltaR is given as a parameter to the estimator so maybe it's that?
fromPV() $>$ 1	I don't see this in the .cc file?
trk$\rightarrow$quality(reco::TrackBase::highPurity)	TrackBase
trk.hitPattern().numberOfValidHits() $\geq$ 8	hitPattern
trk.hitPattern().numberOfValidPixelHits()$\geq$2	hitPattern
trk.normalizedChi2()$<$5	trackBase
std::fabs(trk.dxy(vtx.position()))$<$0.2	trackBase
std::fabs(trk.dz(vtx.position()))$<$17	trackBase

Table 4: Lepton MVA Preselection

Jet distances dxy and dz are measured wrt to the Vertex.

Input Variable	CMSSW Class
lepton pt	pat muon
lepton eta	pat muon
selected track multiplicity of the jet matched to the lepton	see Table 4
rel mini-isolation charged component	lepton.PFIsolation
rel mini-isolation, neutral component	subtract charged miniIso from total miniIso given to MVA (otherwise same class as chargedminiIso)
jet ptRel	calculated from ROOT::XYZTLorentzVector of mu and leading jet
jet ptRatio	pt ratio of ROOT::XYZTLorentzVector of mu and leading jet
CSVv2 b-tagging discriminator of the jet matched to the lepton	JetFloatAssociation::Container
SIP3D	pat muon
lepton dxy	pat muon
lepton dz,	pat muon
muon segment compatibility (only for muons)	pat muon
electron non-triggering MVA ID value (only for electrons)	electrons

4 Low pT MVA (muons only)

This value is calculated using the same class as the Muon MVA so everything is the same but I guess the preselection differs by the following according to the twiki (bolded):

- pt > 5 GeV
- loose PF ID for muons
- **SIP3D < 4**
- **abs(dxy) < 0.5 && abs(dz) < 1**

In PATMuonProducer.cc the estimator object is made from the MuonMVA class, but uses a different input training (weights?) file called "lowPtmvaTrainingFile" instead of "mvaTrainingFile" used for the regular MuonMVA. Like for the Muon MVA, the preselection for loose muons is also done in CMSSW PAT muon producer.