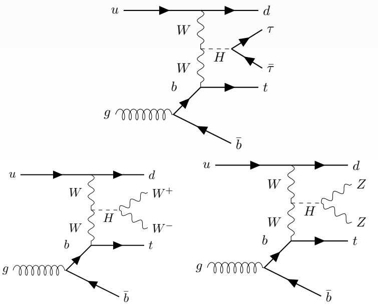
Neural network analysis for hadronic tau channels

Christian Kirfel

9th of December 2020

UNIVERSITÄT BONN

General ML selection



- n-jets: 2 (b-jets: 1)
- b-jet WP: 70 DL1r
- nLeptons & nTaus: 1e/μ 2τ_{had}
- $E_{\mathrm{T,miss}}$: no cut (to 800 GeV)

- jets:
  - $p_T > 35$ GeV
  - $|\eta| < 4.5$
  - EMPFlow
- electrons:
  - $p_T > 20$ GeV leading 27 GeV
  - $|\eta| < 2.5$ not in 1.37 - 1.52
  - WP: LooseAndBLayerLH ; isolation: no requirement
- muons:
  - $p_T > 20$ GeV leading 27 GeV
  - $0.01 < |\eta| < 2.5$
  - WP: Loose ; isolation: no requirement
- taus:
  - $p_T > 20$ GeV leading 27 GeV
  - $|\eta| < 2.5$ not in 1.37 - 1.52
  - WP: RNNLoose
  - ASG recommended OLR (τ_{had} remove jets)

Training specifications

- Training a deep neural network
- Coarse optimisation using an evolutionary neural network
- Fine optimisation doing a grid search
- Optimised hyperparameters:
  - Number of nodes
  - Number of layers
  - Dropout percentage
  - Activation function
  - Weight initialisation
  - Optimiser
- Signal: tHq, tZq
- Background: Diboson, Z+jets, ttbar
- Using absolute weights

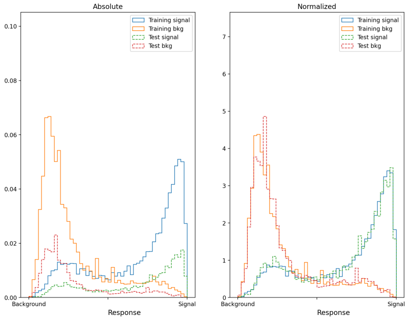
Features

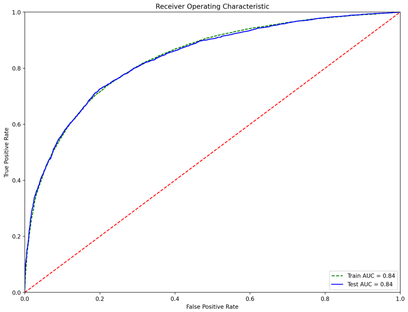
eta_jf	forward jet eta
pt_jf	forward jet transverse momentum
mass_jf	forward jet mass
phi_jf	forward jet phi
eta_b	b-jet eta
pt_b	b-jet transverse momentum
phi_b	b-jet phi
HvisMass	mass of LorentzV sum of hadronic taus
m_met	Missing energy
Reco_w_mass_2	Reconstructed mass of the W case 1
Reco_w_mass_1	Reconstructed mass of the W case 2

deltaRTau	Delta R of the hadronic taus
deltaPhiTau	Delta phi of the hadronic taus
HvisPt	of LorentzV sum of hadronic taus
HvisEta	of LorentzV sum of hadronic taus
TvisMass	mass of reconstructed top
TvisPt	pt of visible top
TvisEta	eta of visible top
M_b_jf	Mass of LorentV sum of b and jf
HT	Sum of transverse energies
lep_Top_pt	Light lepton pt
lep_Top_eta	Light lepton eta

Hyperparameter	Setting
Nodes	180
Layers	3
Dropout	0.8
Batchnormalisation	On
Activation	elu
Output activation	sigmoid
Batch size	1000
Optimisation	Adam
Weight Initialisation	Lecun Normalisation

Results





- Really good separation for this early stage of Optimisation
- Slight problems in training and validation agreement

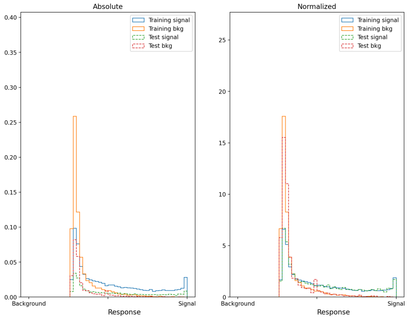
Lorentz invariant variables

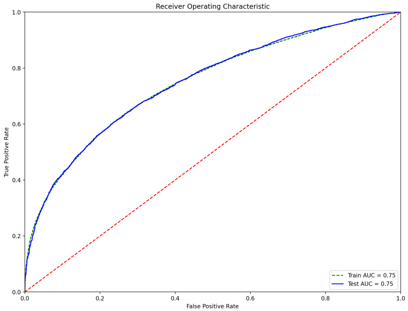
- ip_l1l2
- ip_l1l3
- ip_l2l3
- ip_l1j1
- ip_l1j2
- ip_l1j3
- ip_l1j4
- ip_l1j5
- ip_l2j1
- ip_l2j2
- ip_l3j1
- ip_l3j2

- ip_j1j2
- ip_hadhad
- ip_had1lep
- ip_had2lep
- ip_had1b
- ip_had2b
- ip_lepb
- ip_had1jf
- ip_had2jf
- ip_lepjf
- ip_bjf
- m_met

Hyperparameter	Setting
Nodes	240
Layers	2
Dropout	0.9
Batchnormalisation	On
Activation	elu
Output activation	sigmoid
Batch size	1000
Optimisation	Adam
Weight Initialisation	Lecun Normalisation

Results





- Bad separation
- Only good point is the stability

Lephad Features

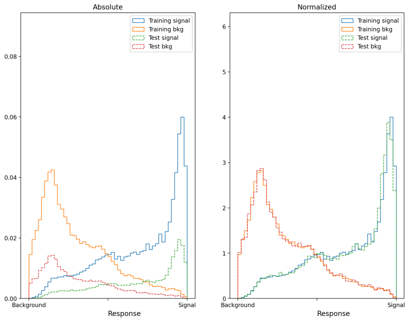
OSDF_LepHad	opposite sign, different flavour
OSSF_LepHad	opposite sign, same flavour
OS_LepHad	opposite sign
OS_ee	opposite sign ee
OS_emu	opposite sign emu
OS_mue	opposite sign mue
OS_mumu	opposite sign mumu
Reco_tautau_mass_1	Reconstructed mass of the taus case 1
Reco_tautau_mass_2	Reconstructed mass of the taus case 2
Reco_w_Tmass_1	Reconstructed transverse mass of the W case 1
Reco_w_Tmass_1	Reconstructed transverse mass of the W case 2

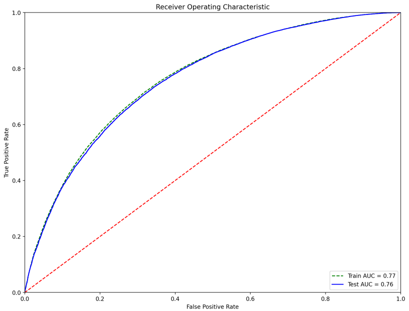
Reco_W_mass_1	Reconstructed mass of the w case 1
Reco_W_mass_2	Reconstructed mass of the w case 2
bscore_jet1	bscore jet1
bscore_jet2	bscore jet2
m_met	Missing energy
vis_tautau_mass	Visible mass of the taus
vis_tautau_pt	Visible pt of the taus
vis_top_mass_lep1	Visible top mass using lepton 1
vis_top_mass_lep2	Visible top mass using lepton 2
vis_top_pt	Visible pt of the top
vis_top_pt_lep1	Visible top pt using lepton 1
vis_top_pt_lep2	Visible top pt using lepton 2

Lephad Hyperparameters

Hyperparameter	Setting
Nodes	180
Layers	3
Dropout	0.8
Batchnormalisation	On
Activation	elu
Output activation	sigmoid
Batch size	1000
Optimisation	Adam
Weight Initialisation	Lecun Normalisation

Lephad Results





- Slightly worse but more stable separation
- Better results expected for additional variables soon to come

Conclusion and future steps

- Both neural network separations show very good separation for an early stage of optimisation
- For the hadhad channel the separation is outstanding but the stability is bad. This means the agreement between training and validation is often bad or has regions that are not understood.
- In both cases the variables have the most significant impact on the result. Different sets of features need to be explored to find better separations.
- A decorrelation and ranking would be desireable
- In the case of lephad several variables have to be added.
- The handling of negative weights is an ongoing problem that needs to be investigated to see if a conclusion can be reached for neural networks.

1. Understand impact of negative weights on variable shape
2. Investigate variable shapes for different treatments
3. Understand impact of negative weights in ML algorithms
4. Test different approaches for negative weight handling and investigate the model stability.

Evolutionary neural networks

- Starting with a set of random configurations
- Evaluate the results of the first generation and generate a new generation based on
- Repeat until a good configuration is reached
- Advantages:
  - Decrease user bias for hyperparameter choice
  - Optimised to run on worker nodes
  - Quick discarding of bad configurations
  - User friendly for unexperienced students

Inner product

$$a^\mu = \begin{pmatrix} p_T \cosh(\eta) \\ p_T \cos(\phi) \\ p_T \sin(\phi) \\ p_T \sinh(\eta) \end{pmatrix}$$

$$\begin{aligned} \langle A|B\rangle &= A_\mu B^\mu \\ &= p_{T,A}p_{T,B}(\cosh(\eta_A)\cosh(\eta_B) - \cos(\phi_A)\cos(\phi_B) \\ &\quad - \sin(\phi_A)\sin(\phi_B) - \sinh(\eta_A)\sinh(\eta_B)) \\ &= p_{T,A}p_{T,B}\left(\cosh(\eta_A - \eta_B) - \cos(\phi_A - \phi_B)\right) \end{aligned}$$