山东大学 SHANDONG UNIVERSITY

Department of Physics, Shandong University

Compressed EWK study(ISRC1N2)

Chengxin Liao
liaocx@ihep.ac.cn

Feb, Mon 24, 2024

Hyperparameters optimization(LH)

Input(LH-Channel):

Sample:

Sig: ISRC1N2(mass_C1 = 100GeV, mass_N2 = 70GeV)->21225 entries
Bkg: 1703476 entries

All input data(C1N2_100_70 and Bkg) already passed pre-selection

```
Signal     -- training events            : 12735
Signal     -- testing events             : 4245
Signal     -- training and testing events: 16980
Background -- training events            : 1022092
Background -- testing events             : 340692
Background -- training and testing events: 1362784
```

Strategy:

method: BDTG
Separate sig(bkg) into five folders, one for test, the other three for train, and last one for validation set,then traverse all possibilities.

Pre-Selection

$$lep\text{-}had\ \ channel: nTaus \geq 1, nLeps \geq 1$$
$$pass\ MET\ trigger;\ MET \geq 200$$
$$1 \leq nBaseJet \leq 8$$
$$b - Veto$$
$$OS$$

Hyperparameters optimization(LH)

Variables(30):

Obj kinematics
- nBase_Jet
- mt_lep
- e_lep(energy of tau2)

Angular correlations
- dPhitt
- dRtt
- dRt1x
- dPhiMin_xj
- dPhiMax_tj

Event kinematics
- Mll(Invariant Mass of tau1 and tau2)
- METsig
- MT2_50
- Mwh(Invariant Mass of tau1 and MET)
- Mwl(Invariant Mass of tau2 and MET)
- MCT(Transverse Mass Squared)
- Proj_j(Projection of pt jet on zeta)
- Proj_tt(Projection of tau1+tau2 on zeta)
- mtx_tau
- Mtx_lep
- ht_tau
- mt_quad_sum
- mt_sum
- frac_MET_tau1
- frac_MET_tau2
- frac_MET_tt
- frac_MET_sqrtHT_40
- frac_jet_tau1
- frac_jet_tau2
- frac_jet_tt
- MT_taumin
- pt_Vframe

High importance at shiyi's feature

Note:

zeta is bisector direction of tau1 and tau2[PhyUtils::bisector(tau1, tau2)]

Hyperparameters optimization(LH)

Grid Search:

Ntrees: 200, 300, 400, 500
Max Depth: 6, 8, 10, 12
MinNodeSize: 1%, 2%, 3%
Learning Rate: 0.01, 0.05, 0.1

Binned significance: $Z = \sqrt{2((s_i + b_i)\log\left(1 + \frac{s_i}{b_i}\right) - s_i)}$

Show top Zn

	Model Name	Binned Significance	Max Zn	Max Zn Bin
12	400_8_1_001	15.6795	4.31391	192
79	400_10_1_001	15.6755	4.26908	192
92	400_12_1_001	15.6890	4.21178	192
77	400_10_2_001	15.3196	4.11376	191
52	500_10_1_001	15.8304	4.11162	194
13	500_12_1_001	15.8210	4.05346	194
120	400_12_1_01	16.0665	4.02939	199
113	500_10_3_001	15.3232	4.02306	192
0	300_12_1_005	16.1734	4.01739	198
139	400_12_1_005	16.2126	4.00753	199
123	500_12_3_001	15.3067	4.00343	192
118	500_12_1_01	16.0441	4.00080	199
24	500_8_1_01	15.9307	3.99007	199
133	500_8_3_001	15.3061	3.97695	192
97	400_6_3_001	15.0010	3.97216	190
26	300_10_1_01	16.0095	3.96339	199
136	300_12_1_01	16.0204	3.94916	199
128	400_8_3_001	14.9962	3.93255	190
107	200_12_1_005	16.0375	3.93002	197
45	400_12_2_001	15.2724	3.92019	191
88	400_12_3_001	14.9991	3.91396	190

Shiyi's result of LH channel

Top Sig

	hy	sig	zn
400_10_2_0.05	15.3225	3.72536
300_11_1_0.05	15.3127	3.87694
500_10_2_0.05	15.3099	3.60778
400_6_1_0.05	15.3075	3.91373
500_8_1_0.05	15.2990	3.58389
400_8_2_0.05	15.2980	3.74427
300_6_1_0.05	15.2929	4.09837
500_8_2_0.05	15.2891	3.63322
200_11_1_0.05	15.2849	3.92924
300_11_2_0.05	15.2804	3.85617
400_11_2_0.05	15.2780	3.68484
300_8_1_0.05	15.2753	3.82506
300_10_1_0.05	15.2733	3.71921
400_11_1_0.05	15.2701	3.60863
500_6_1_0.05	15.2593	3.84429
200_6_1_0.1	15.2559	3.90950
400_12_1_0.05	15.2554	3.58328
500_10_1_0.05	15.2493	3.49410

Top Zn

	hy	sig	zn
200_6_3_0.05	15.0164	4.29022
200_6_1_0.05	15.0755	4.10077
300_6_1_0.05	15.2929	4.09837
200_10_2_0.05	15.1606	4.09228
200_12_2_0.05	15.1803	4.04800
200_8_2_0.05	15.0857	4.01373
200_8_3_0.05	14.9662	4.01324
200_6_2_0.05	14.9743	3.94396
200_11_1_0.05	15.2849	3.92924
300_6_2_0.05	15.1858	3.91508
400_6_1_0.05	15.3075	3.91373
200_6_1_0.1	15.2559	3.90950
200_12_1_0.05	15.2279	3.90380
400_8_1_0.01	14.6829	3.90189
300_11_1_0.05	15.3127	3.87694
200_8_1_0.05	15.1285	3.85623
300_11_2_0.05	15.2804	3.85617
500_6_1_0.05	15.2593	3.84429

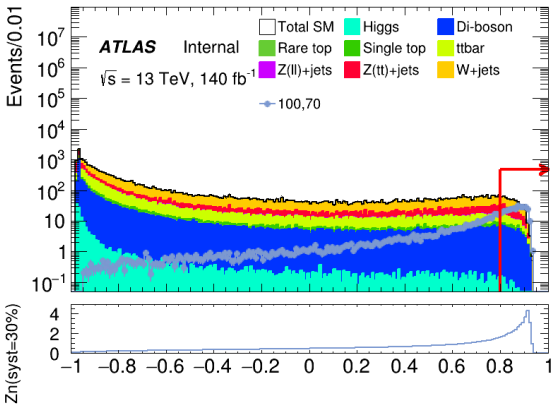
```
Model Name,Binned Significance,Max Zn,Max Zn Bin,Rebin
400_8_1_001,15.6795,4.31391,192,200
400_8_1_001,15.4557,3.67354,96,100
400_8_1_001,15.1841,3.11958,49,50
400_8_1_001,15.309,3.67354,39,40
400_8_1_001,14.74,3.11958,25,25
400_8_1_001,15.0619,3.67354,20,20
400_8_1_001,13.9516,1.66711,10,10
```

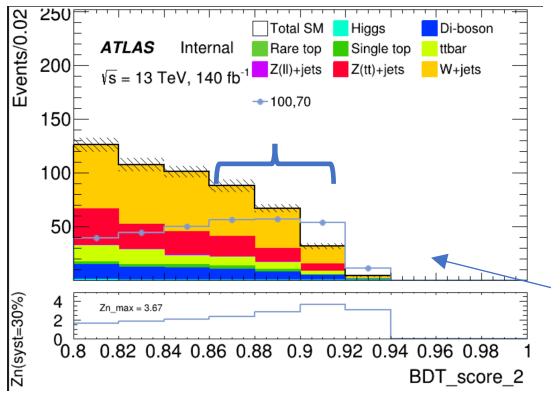
```
400_12_1_01,16.0665,4.02939,199,200
400_12_1_01,15.9754,4.02939,100,100
400_12_1_01,15.6999,3.29277,50,50
400_12_1_01,15.4897,2.90575,40,40
400_12_1_01,15.0231,2.30028,25,25
400_12_1_01,14.7883,2.06337,20,20
400_12_1_01,13.6598,1.39451,10,10
```

```
400_12_1_001,15.689,4.21178,192,200
400_12_1_001,15.4949,3.52564,97,100
400_12_1_001,15.2434,3.52564,49,50
400_12_1_001,15.3089,3.52196,39,40
400_12_1_001,14.8653,3.52564,25,25
400_12_1_001,15.0506,3.52196,20,20
400_12_1_001,13.9276,1.6473,10,10
```

Performance of Model(LH)

hyper parameter: NTrees=400, learning rate=0.01, max depth=8, MinNodeSize=1%(default)





Cut at BDT_score = 0.8, rebin to 100 bins

Notice there is a gap at the edge

3 bins

bin	max Zn	C1N2ISR (100,70)	bkg	Higgs	OtherTop	SingleTop	TopPair	VV	Wjets	Zlljets	Zttjets
94 (0.86-0.88)	2.39932	**55.995+-1.433**	**87.737+-5.078(5.78%)**	0.127+-0.034	0.027+-0.018	2.737+-0.600	8.356+-1.136	10.027+-0.535	**46.913+-5.404**	0.629+-0.128	18.921+-1.943
95 (0.88-0.90)	2.85787	**56.966+-1.438**	**66.750+-5.910(8.85%)**	0.069+-0.023	0.071+-0.029	2.058+-0.482	6.606+-1.013	7.540+-0.405	**37.355+-3.540**	0.619+-0.127	12.433+-1.240
96 (0.90-0.92)	3.67354	**53.908+-1.411**	**31.481+-2.782(8.83%)**	0.021+-0.011	0.025+-0.016	0.621+-0.196	2.791+-0.663	4.435+-0.322	**16.842+-2.246**	0.150+-0.082	6.597+-1.453