山东大学 SHANDONG UNIVERSITY

Department of Physics, Shandong University

Compressed EWK study(ISRC1N2)

Chengxin Liao
liaocx@ihep.ac.cn

Wed, Feb 5, 2024

Outline

1. Hyperparameters optimization
2. Performance of Model

Task-list

- Machine learning for LH channel
  - ~~check more Variable and select significance var for ML(DONE)~~
  - ~~BDTG hyperparameters optimization/ Setup a Grid Search framework (DONE)~~

- Preliminary study on multibody quantum mechanics (In Progress)
  QFT Lecture (Peskin part I)

- BSc thesis: https://www.overleaf.com/project/674e7119837a2580151a0868

- CS61A (python): https://cs61a.vercel.app/index.html

Hyperparameters optimization

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

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.

```
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
```

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

Variables:

Obj kinematics
- pt_lep
- pt_tau
- mt_tau
- e_lep(energy of tau2)

Angular correlations
- dPhit1x
- dRt1x
- dRtt
- dPhitt

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)

- mt_quad_sum
- mt_sum
- frac_MET_tau1
- frac_MET_tau2
- frac_MET_sqrtHT_40
- frac_jet_tau1
- frac_jet_tau2
- frac_jet_tt

Note:

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

Hyperparameters optimization

Grid Search:

Ntrees: 200, 300, 400
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
123	400_12_1_01	14.8848	3.05602	198
137	300_12_1_01	14.8236	2.99004	198
90	400_12_3_001	13.8790	2.98769	189
138	400_10_1_01	14.7405	2.95654	198
46	400_10_3_001	13.8773	2.95293	189
24	400_10_2_01	14.7362	2.94729	197
18	400_8_1_01	14.6892	2.92850	197
140	400_12_1_005	14.8200	2.89803	197
129	400_8_3_001	13.8466	2.89472	189
80	400_10_2_001	14.1171	2.88777	190
48	400_8_2_001	14.0737	2.88315	190
36	300_10_1_01	14.6952	2.86815	197
25	100_10_1_005	14.3038	2.86286	192
112	200_12_1_005	14.6835	2.86017	195
52	400_12_2_001	14.1184	2.85548	190
85	200_12_1_01	14.7532	2.84807	197
0	300_12_1_005	14.7486	2.84744	196
76	100_12_2_005	14.1753	2.84528	191
50	400_12_2_01	14.6361	2.83824	198
6	400_8_1_005	14.6725	2.80702	197
120	300_8_1_005	14.6021	2.80602	196
49	400_10_1_005	14.6962	2.79191	197
101	400_6_3_001	13.7790	2.78598	189
61	300_10_1_005	14.6454	2.78360	196
82	300_12_1_001	14.0489	2.78242	186
143	300_12_2_01	14.5987	2.77344	197
95	400_12_1_001	14.2632	2.74588	191
124	300_10_2_01	14.6525	2.74520	197

Shiyi's result of LH channel

	hy	sig	zn
400_10_2_0.05		15.3225	4.47044
300_11_1_0.05		15.3127	4.65233
500_10_2_0.05		15.3099	4.32933
400_6_1_0.05		15.3075	4.69647
500_8_1_0.05		15.2990	4.30067
400_8_2_0.05		15.2980	4.49312
300_6_1_0.05		15.2929	4.91804
500_8_2_0.05		15.2891	4.35987
200_11_1_0.05		15.2849	4.71509
300_11_2_0.05		15.2804	4.62741
400_11_2_0.05		15.2780	4.42181
300_8_1_0.05		15.2753	4.59008
300_10_1_0.05		15.2733	4.46305
400_11_1_0.05		15.2701	4.33036
500_6_1_0.05		15.2593	4.61315
200_6_1_0.1		15.2559	4.69140
400_12_1_0.05		15.2554	4.29994
500_10_1_0.05		15.2493	4.19292
300_12_1_0.05		15.2461	4.45857
300_6_1_0.1		15.2424	4.48235
300_10_2_0.05		15.2319	4.52470
200_8_1_0.1		15.2283	4.44394
200_12_1_0.05		15.2279	4.68456
300_8_1_0.1		15.2267	4.28422
500_12_2_0.05		15.2216	4.18619
300_12_2_0.05		15.2161	4.43485
400_6_1_0.1		15.2142	4.28123
400_8_1_0.05		15.2139	4.29323

Comment:
compare with shiyi's result, my Max Zn is still quite low and even lower than HH-channel

TODO

1. check more vars and try other method