ALTAS ttH Project for iSTEP

iSTEP 2014
IHEP, Beijing
August 20-29, 2014

Glen Cowan (谷林·科恩）
Physics Department
Royal Holloway, University of London
`g.cowan@rhul.ac.uk`
`www.pp.rhul.ac.uk/~cowan`

Introduction

Goal: search for the production of a Higgs in conjunction with a ttbar pair in ATLAS ("ttH") with H → bb.

MC training data in *n*-tuples `tth_sig.root, tth_bkg.root`

      Signal: ttH Pythia8, m_H = 125 GeV

      Background: inclusive tt from Powheg + Pythia8

Preselection: Required 4 b-tagged jets + 2 opposite sign leptons

Effective luminosity after preselection:

      Signal: $L_{MC} = 6.40 \times 10^4$ fb^{-1}

      Background: $L_{MC} = 1.09 \times 10^2$ fb^{-1}

Find two b-jets with mass closest to 125 GeV

      → Higgs candidate

Associate the other two b-jets with a lepton by lowest mass first

      → top candidates (missing the neutrinos)

Variables for MVA (updated)

TTree n-tuple **mva** contains for each event:

w	event weght
mtot	total mass of 4jet + 2lepton system
met	missing transverse energy
m34	mass of Higgs candidate
pt34	p_T of Higgs candidate
ylo,yhi	$y = m^2(\text{top-Higgs}) / m^2_{\text{tot}}$, order: $y_{\text{hi}} > y_{\text{lo}}$
ctbz	cos (angle between Higgs plane and ttbar plane)
cts	cos (decay angle of b-jets for Higgs candidate)
ctll	cos (angle between leptons)
pvxp_n	number of pile-up vertices in event

Naive counting analysis

Find expected number of events for Ldata = 20.5 fb-1

   = sum of weights for selected events $\times L_{\text{data}} / L_{\text{MC}}$

Find for expected numbers of events for both signal and background with no further cuts, s_{tot} and b_{tot}.

Naive median discovery significance

$$\text{med}[Z|\text{s}] = s_{\text{tot}} / \sqrt{b_{\text{tot}}}$$

Goal is to use MVA to construct test statistic.

Cut on test statistic; expected numbers of events after cut:

$$s = s_{\text{tot}} \; \varepsilon_s = s_{\text{tot}} P(t > t_{\text{cut}}|\text{s})$$
$$b = b_{\text{tot}} \; \varepsilon_b = b_{\text{tot}} P(t > t_{\text{cut}}|\text{b})$$

Goal is to maximize the discovery significance after cut.

Code

TMVA Code is in:

`www.pp.rhul.ac.uk/~cowan/stat/beijing14/`

`istep_atlas_project.tar`

Download to working directory and type

`tar -xvf istep_atlas_project.tar`

YOUR PROJECT:
1) Find s_{tot} and b_{tot} for $L_{data} = 20.5$ fb^{-1}
2) Train coefficients of Fisher discriminant (`tmvaTrain`)
3) Use cut on Fisher disc. to get s and b versus t_{cut}. (`analyzeData`)
4) Plot $s/\sqrt{b}$ versus t_{cut} and find maximum (use `TGraph`)
5) Find L_{data} required to have $s/\sqrt{b} = 5$.
6) Repeat for MLP and BDT (see TMVA manual).