Optimization of sensitivity for the Di-Higgs search in $b\bar{b}l\nu qq$ final state in the LHC data recorded by the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV

Pratik Kafle *

Advisor: Dr. Suyog Shrestha†

November 22, 2020

Abstract

A study of optimization of signal sensitivity in the search for the diHiggs event in the ATLAS experiment is presented where one of the Higgs bosons decays via the $H \to b\bar{b}$ channel and the other via $H \to WW^*$ with $b\bar{b}l\nu qq$ as the final state . A large irreducible background contamination from $t\bar{t}$ decay significantly decreases the signal sensitivity in the diHiggs search. We employ ROOT's built-in Multivariate analysis tool(TMVA) to train new physics motivated variables along with currently trained variables. An increment in the rejection of $t\bar{t}$ background while keeping the signal efficiency high is observed.

1 Introduction

In Standard model(SM), elementary particles acquire mass through the Electroweak Symmetry Breaking Mechanism(EWSB). The EWSB mechanism is strongly suggested by the discovery of a SM Higgs boson-like particle of mass near 125 GeV by the ATLAS and CMS collaboration in 2012 [1, 2]. This discovery has opened the door for the close study the properties of the Higgs Boson with greater precision. Close study of the properties of the Higgs Boson will either help strengthen the Standard model or pave a path to find physics beyond the Standard model. One of the ways to probe properties of Higgs boson is to study Higgs self–coupling. The scalar potential of the Higgs doublet field ϕ that is responsible for spontaneous electroweak symmetry breaking, which can be reinstated only by studying the Higgs self interaction as given by [3]:

$$V_H = \mu^2\phi^\dagger\phi \, + \, \frac{1}{2}\lambda(\phi^\dagger\phi)^2 \, ; \qquad \lambda = \frac{{M_H}^2}{\nu^2} \text{ and } \nu^2 = -\frac{1}{2}M_H^2, \tag{1}$$

*Department of Physics, Reed College

†Ohio State University

S.N	Production Mode	Cross Section(fb)
1	ggF-hh	40
2	VVF-hh	2
3	V-hh	1
4	tt-hh	1

Table 1: Production cross section of diHiggs via different mechanism at the LHC at $\sqrt{s}$= 14 TeV.

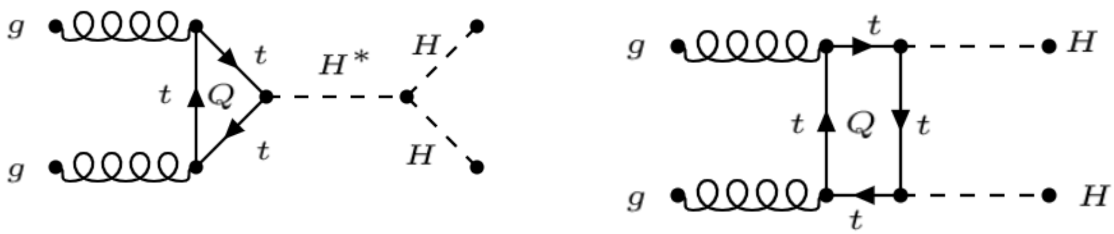


Figure 1: The leading order feynmann diagrams showing the production of SM di-Higgs through gluon-gluon fusion. The Figure in the left has Higgs self-coupling, where a non-resonant Higgs particle is formed. However, the figure in the right is a simple production of two resonant Higgs particles.

In these equations, λ is the Higgs self-coupling parameter and μ is a constant. The vacuum expectation value, ν, is related to λ and μ by the equation, $\nu = \frac{|\mu|}{\sqrt{\lambda}} = \frac{M_W}{g}$ is 246 GeV. In SM, the trilinear Higgs self-coupling can be related to the the mass of Higgs Boson(M_H) as:

$$\lambda_{HHH} = \frac{3M_H^2}{\nu} \tag{2}$$

There are four suggested ways of production of diHiggs at the Large Hadron Collidor(J. Baglio, A. Djouadi, R. Grober, M. M. Muhlleitner, J. Quevillon, M. Spira, 2012), namely:

1. The gluon fusion mechanism, $gg \rightarrow HH$ [4–7], the feynmann diagrams of which are shown in Figure 1,
2. the WW/ZZ fusion processes (VBF), $qq \rightarrow V^*V^*qq \rightarrow HHqq(V = W, Z)$ [8–10],
3. The double Higgs–strahlung process, $q\bar{q}' \rightarrow V^* \rightarrow VHH(V = W, Z)$ [11] and,
4. associated production of two Higgs bosons with a top quark pair, $pp \rightarrow t\bar{t}HH$ [12]

The cross section of production of diHiggs(14 TeV) via each of the above shown methods is shown in table 1 [3].

In this paper, we present a study to optimize the search for the diHiggs decay produced from gluon-gluon fusion with $b\bar{b}ql\nu$ final state in the ATLAS experiment at the LHC using

Monte Carlo simulated sample. A big challenge in observing this event is due to an enormous background contamination, the dominant of which is due to the decay of $t\bar{t}$ decay. We use ROOT and its built-in multivariate analysis toolkit (TMVA) to run a sensitivity estimate using the Boosted Decision Tree algorithm to discriminate signal from the $t\bar{t}$ background. The performance of the algorithm to discriminate the background from signal is based on the Signal significance, which is defined as $S/\sqrt{S+B}$.

2 ATLAS Detector

The ATLAS detector is a general-purpose multi-layered, forward–backward symmetric cylindrical particle detector at the LHC designed to test the existing theories in Standard Model as well as probe physics beyond it. The cylindrical detector is 46m long, 25m in diameter with a weight $\sim$ 7000 tonnes, which makes it the largest volume constructed particle collider [13]. The ATLAS detector consists of four major components: the Inner Detector, the Calorimeter, the Muon Spectrometer and the Magnet System [14]. The Inner Detector, contained in an axial magnetic field, is highly sensitive and compact component built to measure the direction, momentum, and charge of electrically-charged particles produced from the proton-proton collision in the range of $|\eta| < 2.5$ [15,19]. Similarly, Calorimeters are made of high-density material to absorb, and hence detect the energy deposited by particles that are produced from the collisions over the range of $|\eta| < 4.9$ [19][1].

Electromagnetic calorimeters sample the energy of electromagnetically interacting particles, such as electrons and photons as they interact with matter and the Hadronic calorimeters measure the energy deposited by the hadrons. While the Calorimeters can stop most known particles, muons and neutrinos pass through it with a very little interaction [18]. The muons are detected and their properties measured by the Muon spectrometer, which are located outside of the calorimeters. It consists of several highly precise detectors that can track the muons in the range of $|\eta| < 2.4$ and $|\eta| < 2.7$ respectively [19]. Since the detector is bombarded by cosmic muons all the time, the muon tracks are matched to see if they originate from the interaction vertex; this helps identify the muons produced from the collision for analysis.

3 Signal event: Non-resonant di-Higgs decay

The di-Higgs can decay in different ways each with different probabilities. In this paper, we study the case where a non-resonant SM diHiggs decays semi-leptonically with single lepton

[1]

A right handed co-ordinate system is adapted to locate positions in the ATLAS detector, with its origin at the nominal interaction point(IP), which is also the geometric center of the detector. X-axis is the line passing through the IP and the center of the LHC ring, while y-axis points vertically upward. The z-axis is then along the beam pipe. A spherical co-ordinate system (R, ϕ, θ) is also used to describe the LHC, where ϕ is the angle in the transverse plane(x-y) measured from positive x-axis while θ is the polar angle in x-z plane measured form positive z-axis. The pseudorapidity (η) is defined in terms of the polar angle θ as $\eta = \ \ln[\tan(\theta/2)]$. The angular distance between particles and jets in (θ, ϕ) plane is measured by $\Delta R := \sqrt{\Delta\theta^2 + \Delta\phi^2}$

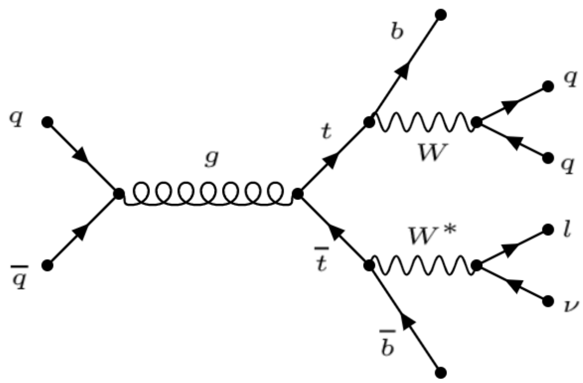


Figure 2: Feynmann diagram showing $t\bar{t}$ decay

in the final state. Equation 3 and the feynmann shown in Figure 3 shows this process.

$$H^* \rightarrow HH \rightarrow b\bar{b}WW^* \rightarrow b\bar{b}ql\nu \tag{3}$$

4 Background event

In the LHC, there are a many background events that give rise to the same final state that of signal events($b\bar{b}ql\nu$). The background contributions come from $t\bar{t}$, single top, diboson, multijet, W/Z+jets, ggH and ttH. Out of these, the dominant is $t\bar{t}$ decay. For simplicity in our analysis, we only consider the background from $t\bar{t}$ decay. The mechanism of $t\bar{t}$ decay contributing to the signal contamination is shown in equation 4, and corresponding feynmann diagram is shown in Figure 2.

$$t\bar{t} \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}qql\nu \tag{4}$$

A luminosity of 140 fb^{-1} is assumed for the analysis of this data, which is recorded either by single electron, MET or single muon triggers. Similarly, a signal cross section of 33 fb and a background cross section of 8000 fb is used in this analysis. While the background cross section is not close to the real data, it can be used to cap the maximum significance in real data. In general, maximum significance is negatively co-relates with the number of background events, thus, this analysis gives a gauge of what we can expect from the real data. Number of expected events is the product of cross section and Luminosity, as shown in Equation 5,

$$N = \sigma \times \mathcal{L} \tag{5}$$

Based on 5, the expected number of signal events is 4620 and that of background is 1120000.

5 Monte Carlo samples

Monte-Carlo(MC) simulation makes use of different computational algorithms to replicate events based on repeated random sampling [16]. For our analysis, the $t\bar{t}$ (nominal) sample,

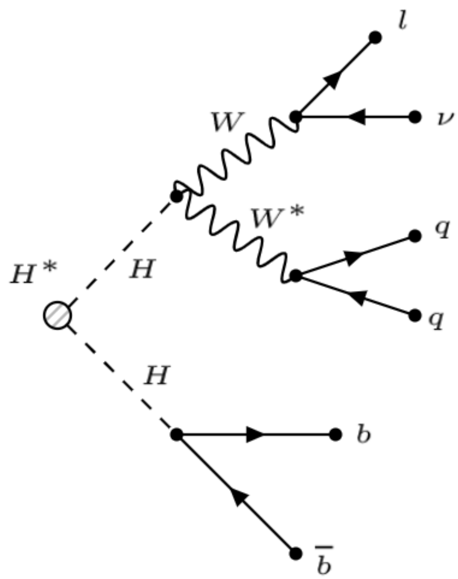


Figure 3: Feynmann diagram of semileptonic decay of Higgs pair with only one lepton in final state.

(mc16_13TeV.410470.PhPy8EG_A14_ ttbar_hdamp258p75_nonallhad) have been simulated using Powheg [26] as generator, Pythia8 (v8.230) [25] for the showering and A14 tune for the hadronisation. The mass of the top quark is set to $m_T = 172.5$ GeV and at least one top quark in the ttbar event is required to decay to a final state with a lepton. EvtGen (v1.6.0) is used for properties of the bottomed and charmed hadron decays. The Hdamp parameter, used to regulate the high-pT radiation in Powheg is set to 1.5 times the mass of the top quark for good data/MC agreement in the high-pT region.

The signal sample (mc16_13TeV.450579.aMcAtNloHerwig7EvtGen_H7UEMMHT_CT10 ME_hh_bbWW) is the non-resonance HH production via gluon-gluon fusion where on Higgs boson subsequently decays into a two b-quarks while the other Higgs boson decays into two W bosons. Events are first generated at Next-to-Leading-Order with MadGraph5_aMCatNLO using the FTApprox method. The CT10 parton distribution function is used in the event generation, which is interfaced to Herwig 7.0.4 using the H7-UE-MMHT tune for underlying events and the H7-MMHT2014LO tune for parton shower and hadronisation. Events are generated with an effective Lagrangian in the infinite top-quark mass approximation, and reweighting the generated events with form factors that take into account the finite mass of the top quark. This procedure partially accounts for the finite top-quark mass effects. The final state of the sample is semi-leptonic, where one of the W bosons is required to decay into a charged lepton and a neutrino, while the other into two quarks.

6 Multivariate Analysis

6.1 Boosted Decision Tree(BDT)

A decision tree(DT) is a binary structured classifier which consists of series of simple conditional control statements to get an outcome [20]. A decision tree is initiated with a root node and then each consecutive layers are formed by binary nodes. A node is split based on a cut applied on a variable, and the process continued recursively. To make a cut in a new node, a variable and cut value which yields the maximum signal and background separation is selected [21]. As the tree grows, the phase space is divided into sub regions until an event is classified as signal or background, depending the bottom(leaf) node. A shortcoming of the Decision trees is the instability with respect to statistical fluctuations in the training sample from which the tree structure is derived [21]. to overcome the limitation, many decision trees are trained together and integrated to form a "forest". A single decision tree in a BDT training can be seen in the Figure 5.

The generation of new trees designates more emphasis to previously misclassified events. Several weak classifiers are amalgamated to construct a strong classifier which has a higher likelihood of correct categorization. This process is called "boosting". After the classification as signal/background, the BDT score, a weighted average of the individual tree classification, is given. This BDT score shows the likelihood of an event being signal or background.

6.2 Receiver Operating Characteristic (ROC) curve

Receiver Operating Characteristic (ROC) curve is one of the tool used to evaluate the performance in the training. It is a plot of background rejection vs Signal Efficiency(Figure 8) for cuts made on BDT output score (figure **??**). In this analysis, Signal efficiency is defined as the proportion of signal retained for a particular cut on the BDT score [24]. Similarly, Background rejection is defined as the proportion of background rejected by that same cut.

Performance of a method can be judged using the convexity of the ROC curve. If a method performs better, then there will be higher background rejection for a particular signal efficiency, meaning a more convex curve. We use ROC to benchmark the variable list performance in this investigation, but other performance benchmarks are possible.

6.3 Overtraining

BDT training faces a major burden of Overfitting. Instead of describing the data set in general, over training starts to outline minute statistical fluctuations in the data set, resulting in a low predictive model. Some of the ways to mitigate this is by training using more training samples, and use cross validation, that is generate multiple mini train-test splits to tune the model.

6.4 Dimension reduction

The quality of a machine learning algorithm declines with the increase in the number of input variables. The difficulty in analyzing and organizing data in high-dimensional spaces that do

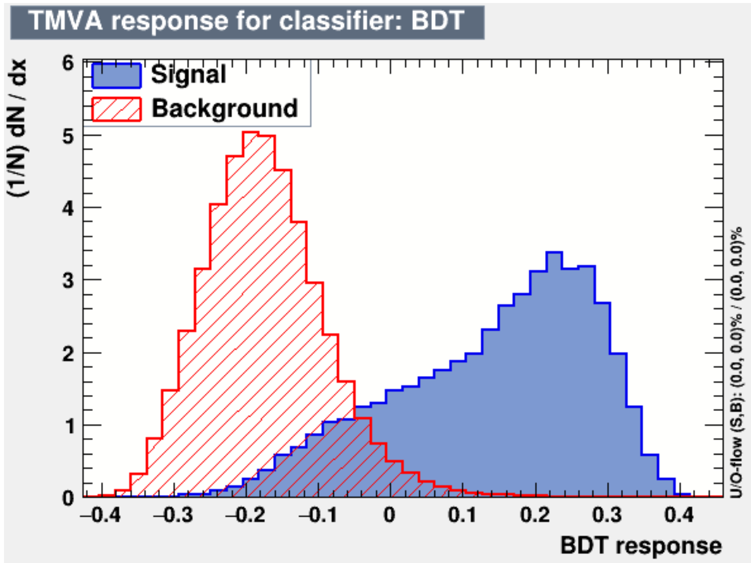


Figure 4: BDT Output score for 43 trained variables. Some proportion of all signal events get selected in addition with some background contamination when a cut is made. "Signal efficiency" is defined as the ratio of selected signal events to total signal events for that cut, and the ratio of rejected background events to total background events is defined as "background rejection".

not occur in low-dimensional setting is often referred as the "Curse of dimensionality" [22]. There are many ways to reduce dimension of the training, such as feature selection and Linear discriminant analysis (LDA) [23]. Co-relation Matrix for Signal and Background (Figure 9 and 10) can be used to eliminate highly co-related variables to reduce the dimension.

7 Results

We used the builtin BDT method from TMVA [21] to train 43 kinematic variables, whose distributions are depicted in Appendix A. In previous analysis, 36 kinematic variables were trained using built in Deep neural network(DNN) and significance evaluated. For our present analysis, we trained 7 variables(emphasised variables in Table 2) in addition to 36 kinematic variables and performed the same analysis. The newly added kinematic variables ranked higher than most of the previously trained variables as shown in Table 3, with **loglratio** ranked as the most important variable. The maximum significance of the new set of BDT trained variables exceeded the previous one by 1.41% for Signal:Background = 4620:1120000. We also performed the same analysis as above and trained the two sets of variables with Deep Neural Network(DNN). As with BDT method, the maximum significance while training with 43 variables was greater than training 36 variables, which is depicted in Figures 6 and 7.

In previous analysis, 36 variables were trained using DNN which gave an output distribution(dHH), which is depicted in Figure 16. The variable dHH has a maximum significance of 35.0064 at Signal:Background = 4620:1120000 as shown in Figure 13 while training dHH with 7 added variables with BDT resulted to a maximum significance of 50.7840 at the same

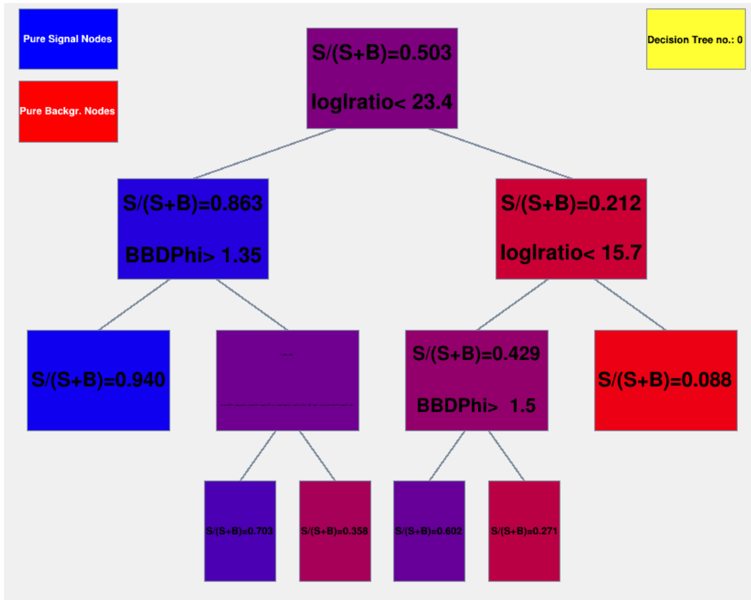


Figure 5: Figure showing the branching based on feature cuts when 43 variables are trained. The red a node is, the more likely it is a signal, and the blue a node is, the more likely it is to be the background.

expected S:B as previous. This is illustrated in Figure 14.

8 Conclusions

The newly added kinematic variables rank higher than the previously trained variables. A 19.4 % increment in the number of new variables led to 1.41% increment in the Signal significance when new kinematic variables are trained with BDT. Similarly, training the output of DNN trained variable, dHH with 7 new variables increased the significance by 34.5%. This is a promising result in terms of increment in signal efficiency. Most of the present analysis is done by training new variables through trial and error. It is promising to work on automating the search for combination of different Kinematic variables to enhance the optimization of the diHiggs search.

Acknowledgments

I'd like to thank Dr. Suyog Shrestha for supervising my project. Similarly, I'd like to thank Center of Life Beyond Reed(CLBR) for funding this research.

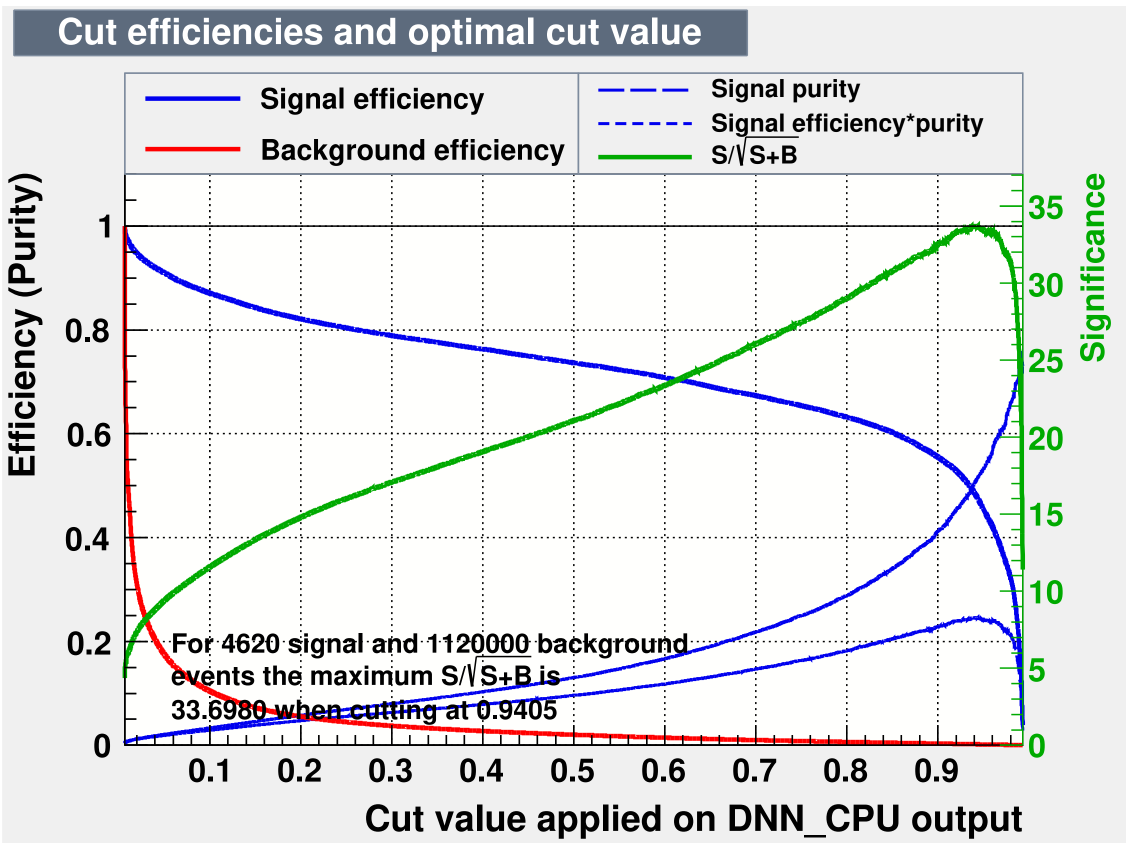


Figure 6: Signal Significance peaked at 33.6980 for 36 DNN trained variables for 4620 signal and 1120000 background events.

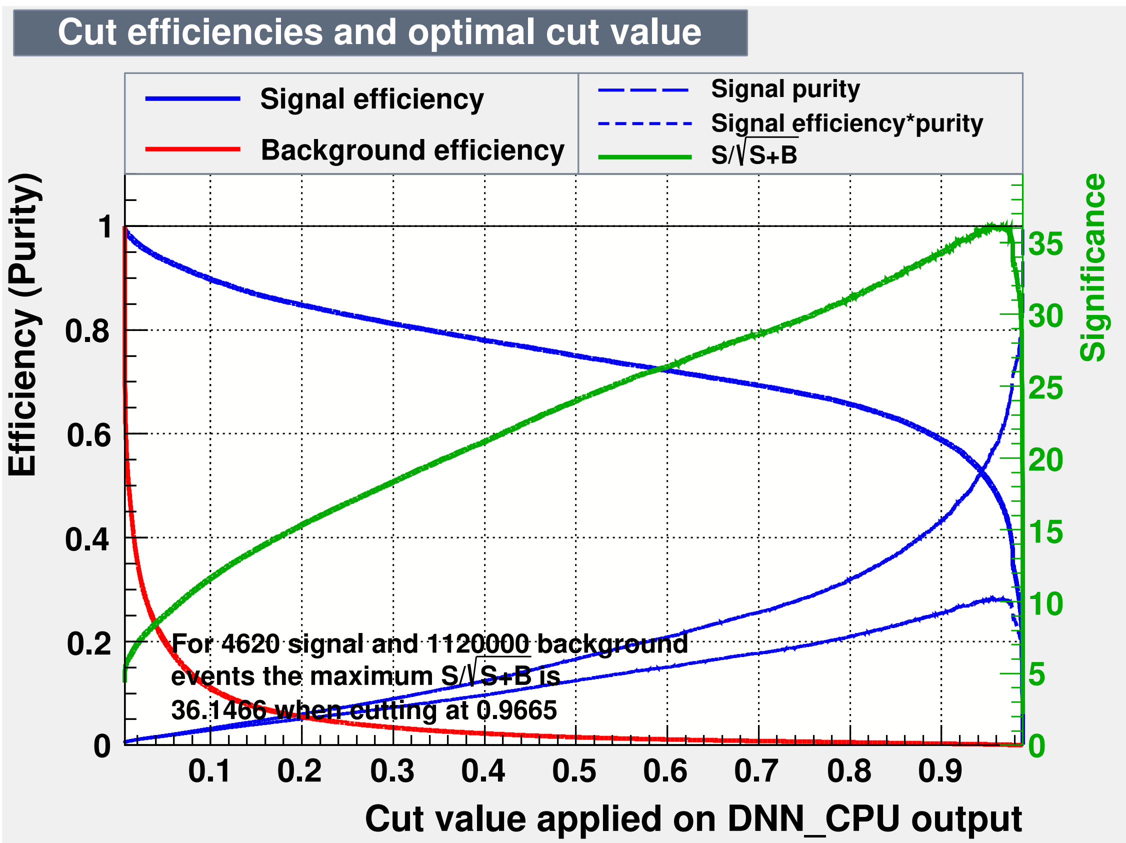


Figure 7: Signal significance peaked at 36.1466 for 43 DNN trained variables for 4620 signal and 1120000 background events.

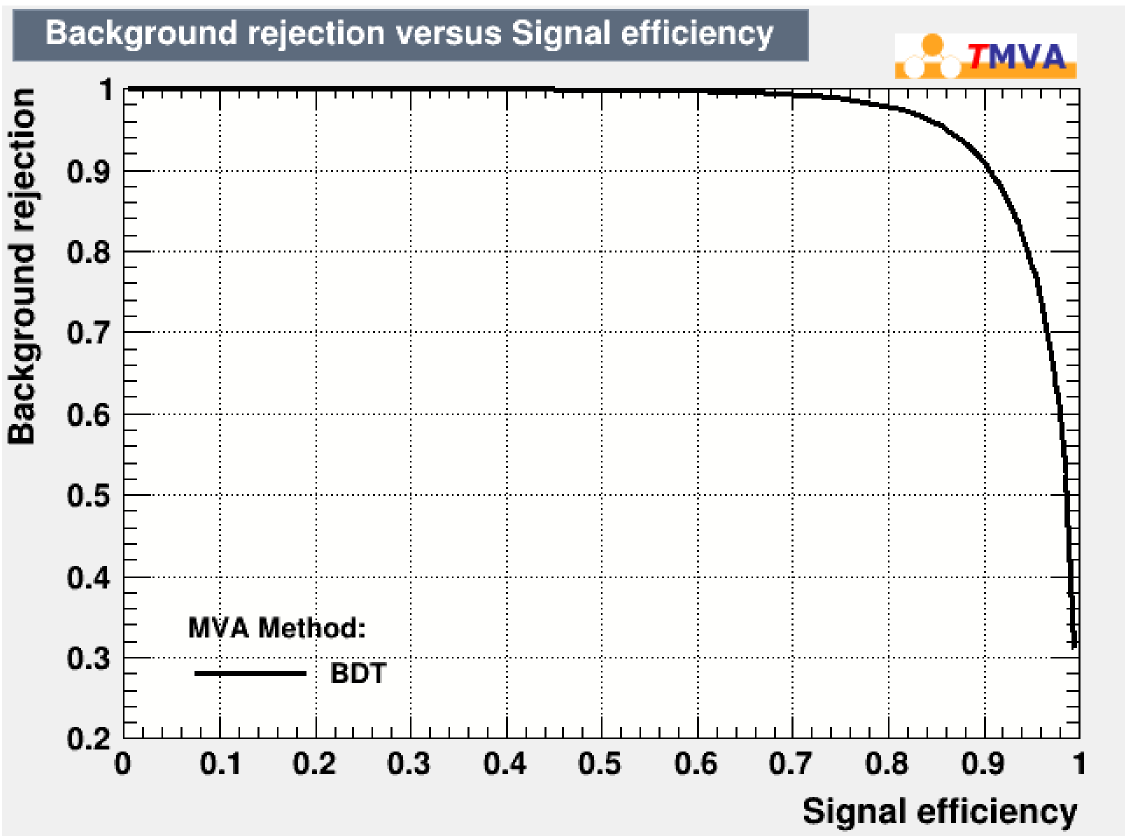


Figure 8: Background rejection vs Signal efficiency(ROC) curve using BDT method while training 43 variables.

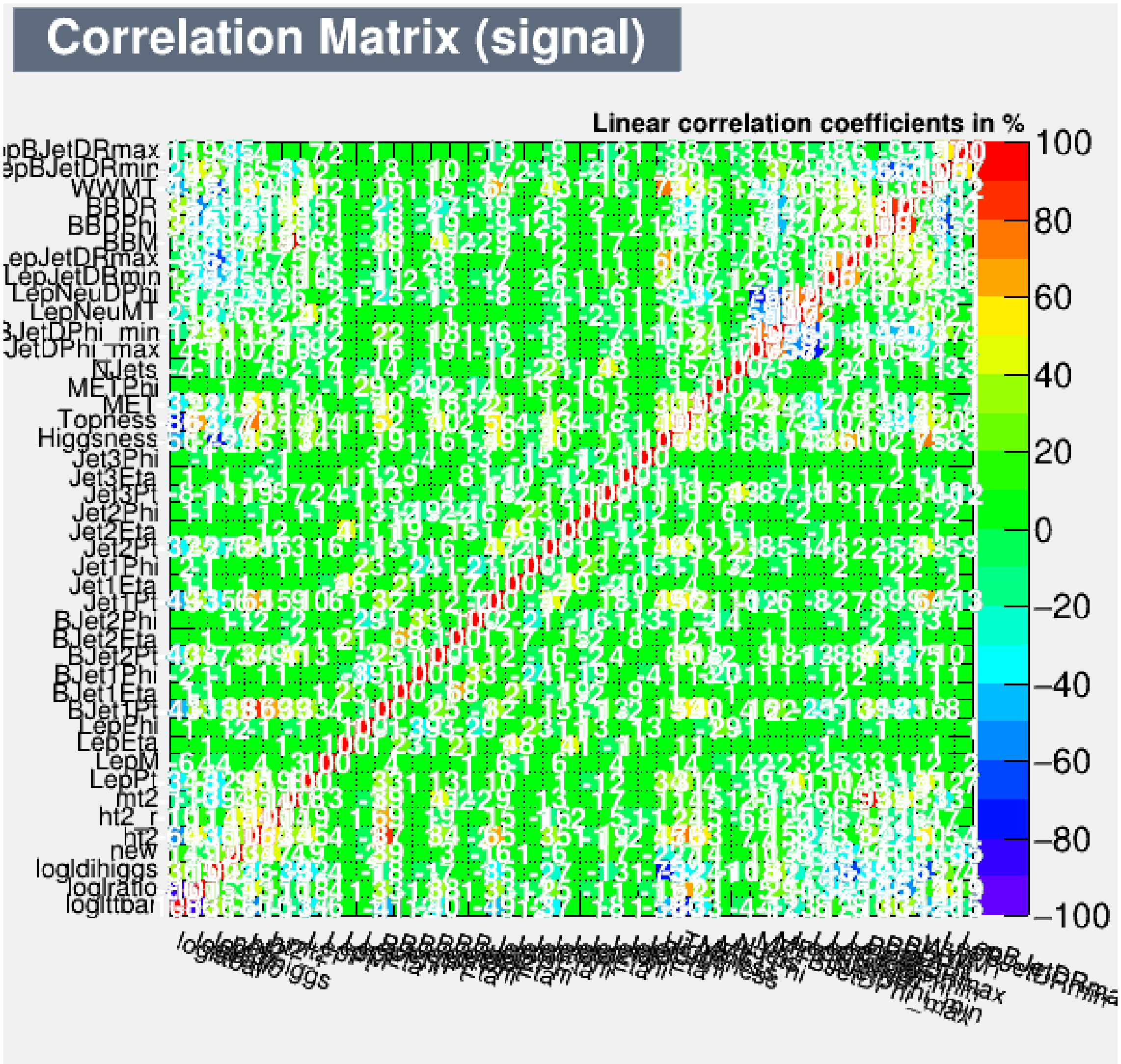


Figure 9: Co-relation Matrix(Signal) for for 43 variables

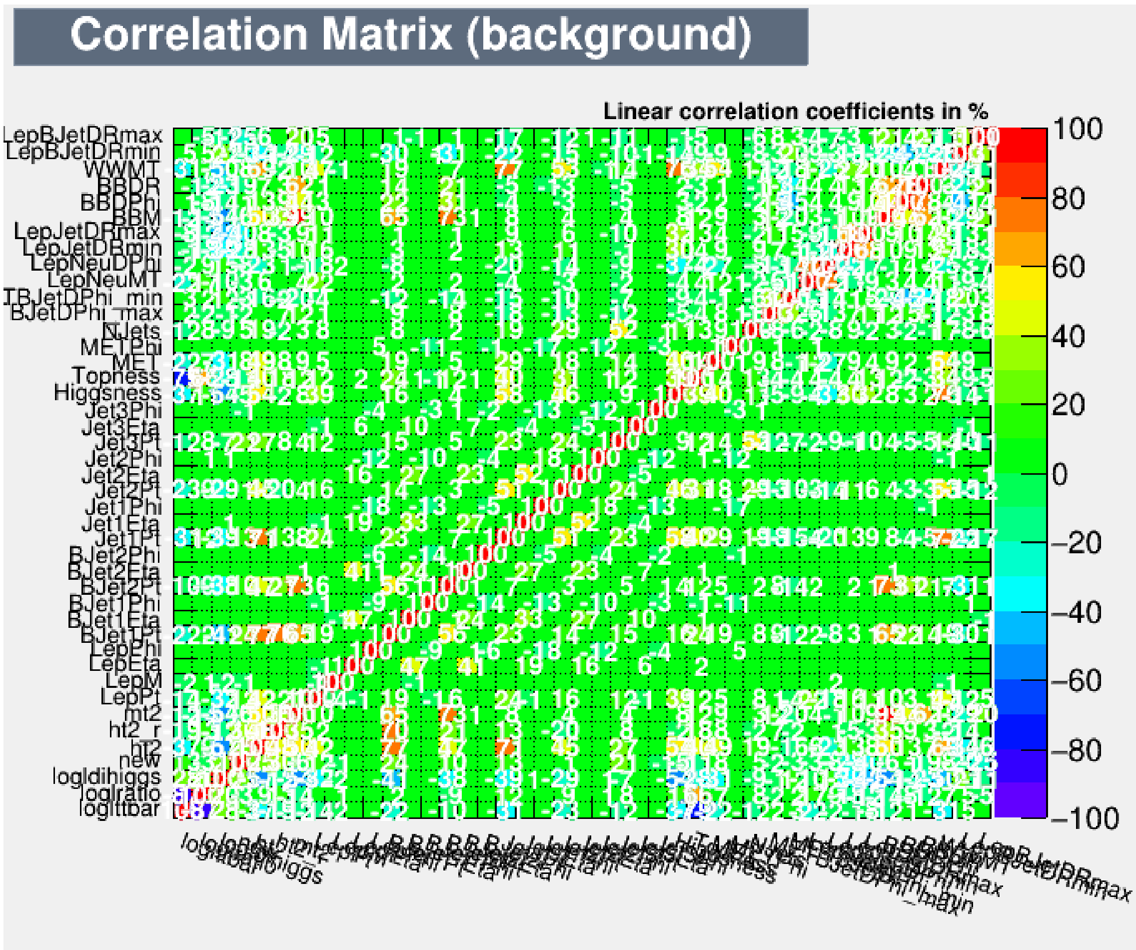


Figure 10: Co-relation Matrix(Background) for for 43 variables

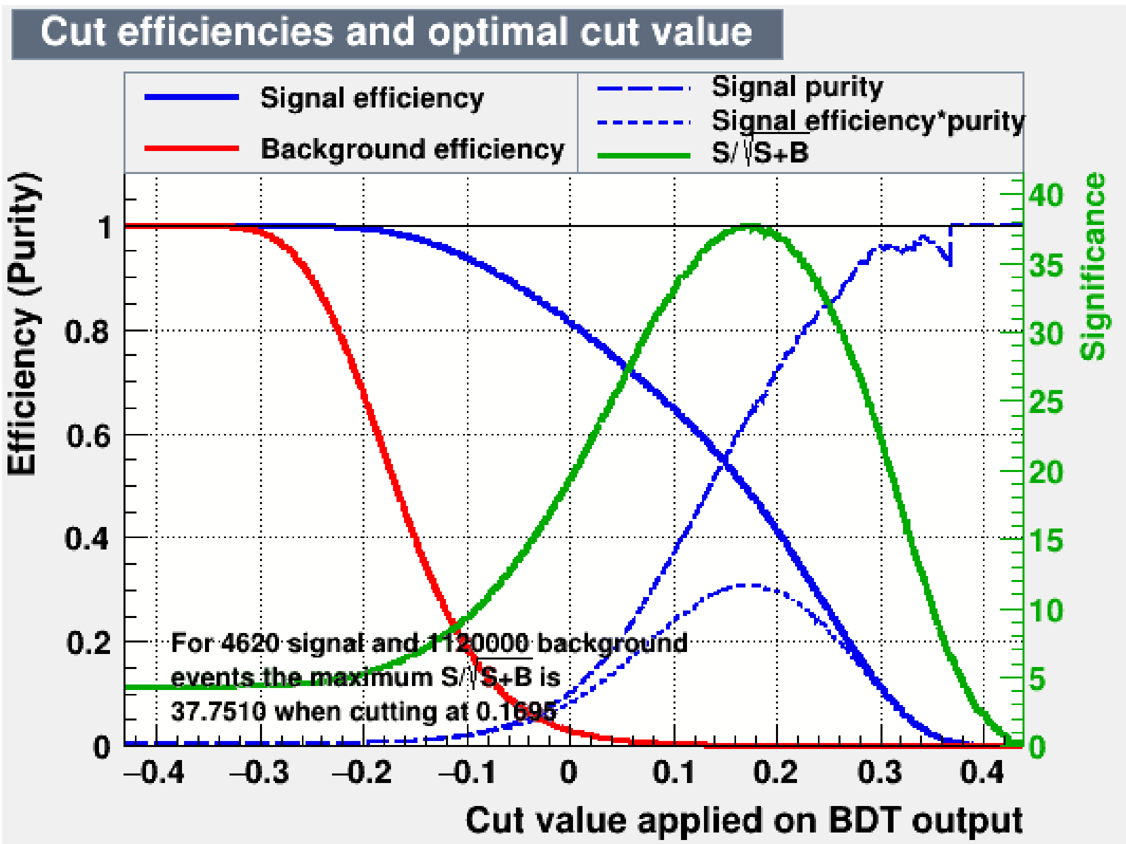


Figure 11: Signal Significance peaked at 37.7510 for 36 BDT trained variables when for 4620 signal and 1120000 background events.

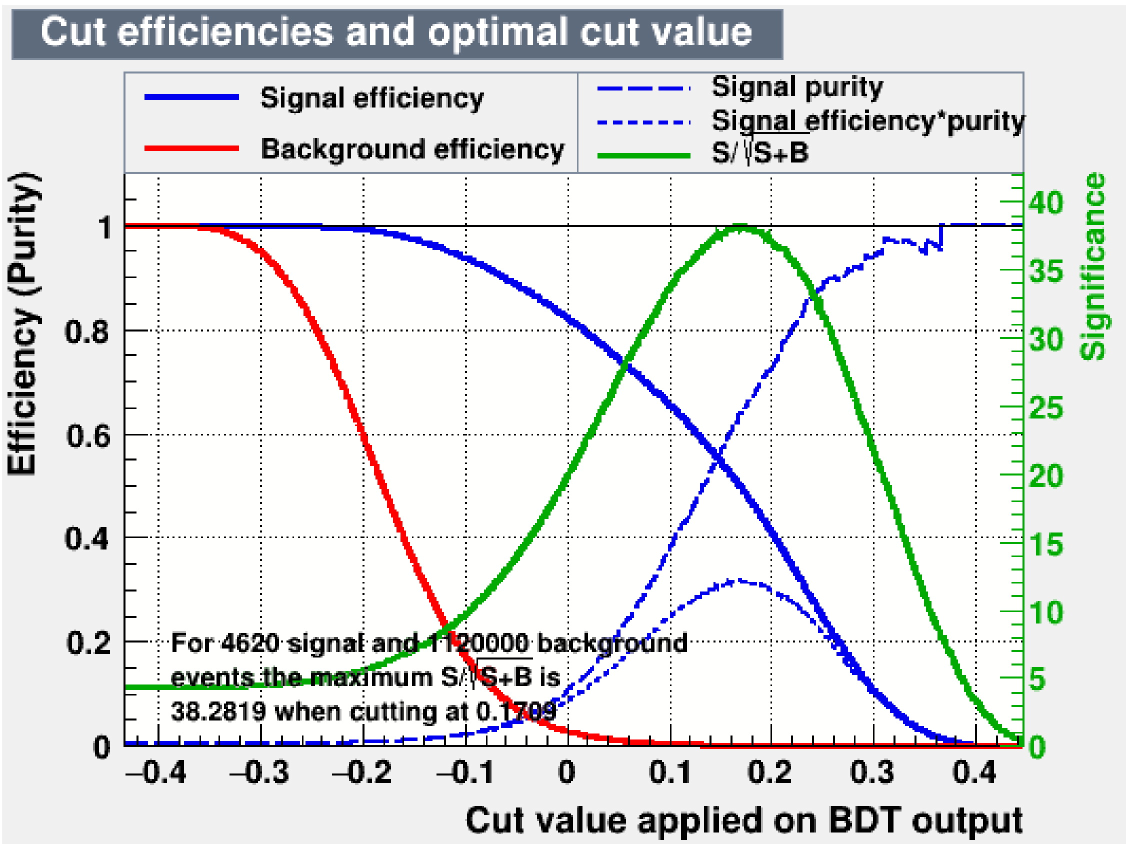


Figure 12: Signal significance peaked at 38.2819 for 43 BDT trained variables when for 4620 signal and 1120000 background events.

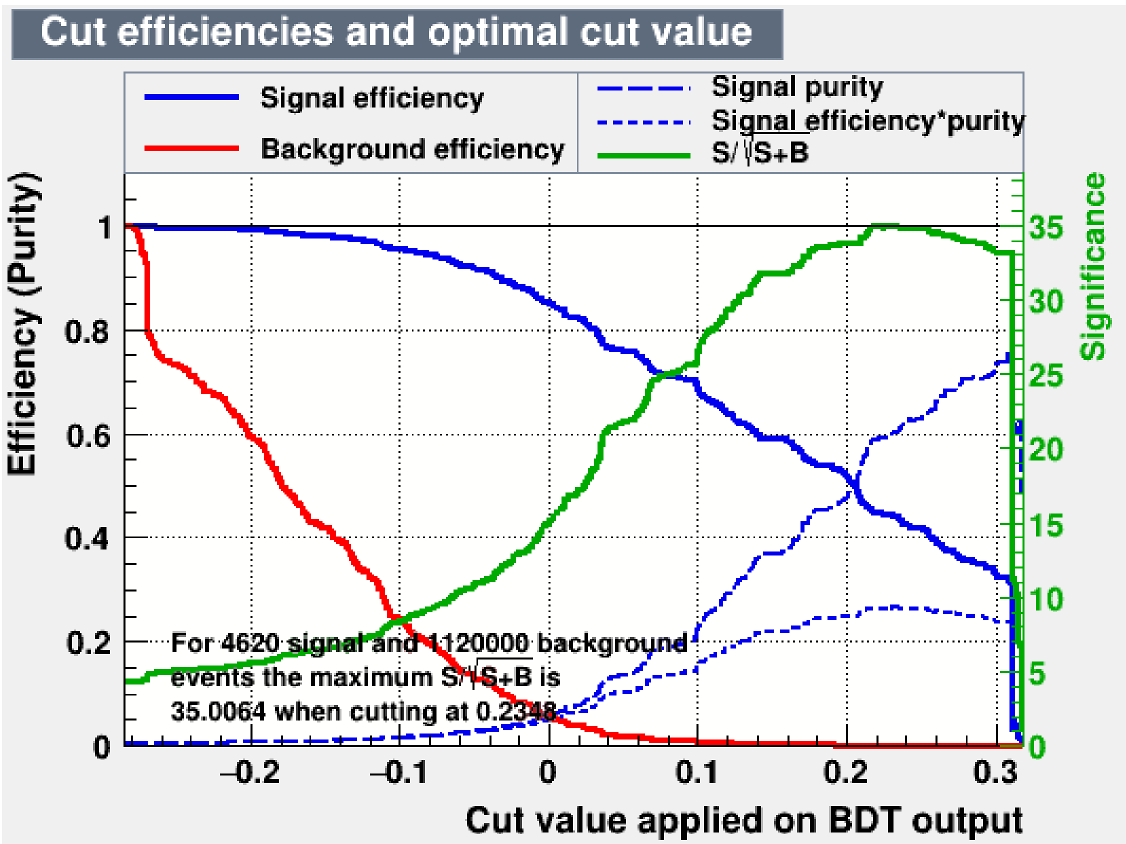


Figure 13: Signal significance peaked at 35.00 when dHH,dHH are BDT trained for 4620 signal and 1120000 background events.

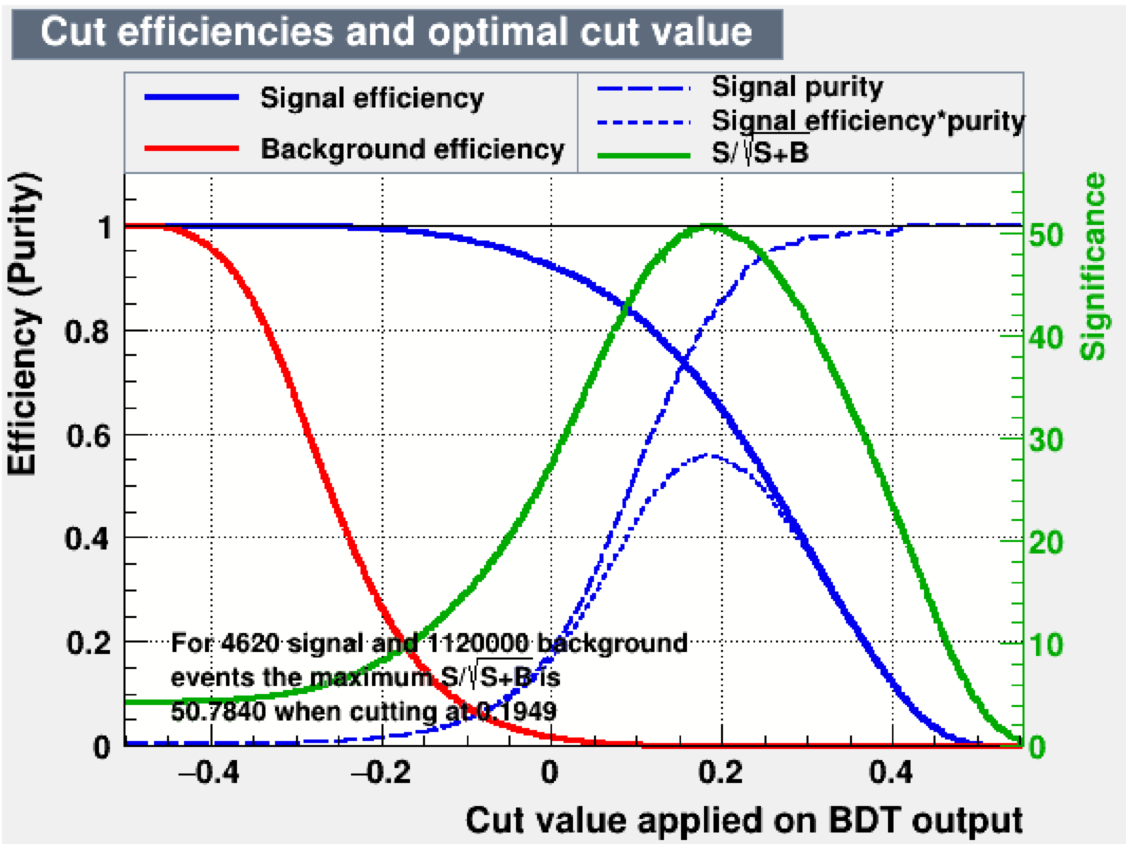


Figure 14: Signal significance peaked at 38.2819 when dHH and 7 variables are BDT trained for 4620 signal and 1120000 background events.

S.N	Var	Sep
1	**loglratio**	5.105e-01
2	**logltbar**	3.916e-01
3	BBDR	3.809e-01
4	Topness	3.538e-01
5	BJet1Pt	3.454e-01
6	LepBJetDRmin	3.437e-01
7	**ht2**	3.317e-01
8	BBDPhi	2.896e-01
9	LepJetDRmin	2.850e-01
10	BBM	2.472e-01
11	**ht2_r**	2.460e-01
12	**mt2**	2.440e-01
13	LepJetDRmax	1.785e-01
14	METBJetDPhi_min	1.778e-01
15	LepBJetDRmax	1.363e-01
16	**logldihiggs**	1.268e-01
17	LepPt	1.056e-01
18	**new**	1.054e-01
19	BJet2Pt	1.035e-01
20	Jet1Pt	8.707e-02
21	METBJetDPhi_max	8.677e-02
22	LepNeuDPhi	8.021e-02
23	WWMT	7.309e-02
24	MET	5.345e-02
25	LepNeuMT	5.201e-02
26	Higgsness	4.691e-02
27	Jet3Pt	3.144e-02
28	Jet2Pt	2.297e-02
29	NJets	1.147e-02
30	Jet3Eta	8.051e-03
31	Jet3Phi	7.821e-03
32	BJet1Eta	7.221e-03
33	BJet2Eta	5.377e-03
34	LepM	2.520e-03
35	BJet1Phi	2.336e-03
36	LepEta	2.259e-03
37	Jet2Phi	2.171e-03
38	METPhi	1.942e-03
39	Jet1Eta	1.895e-03
40	Jet2Eta	1.854e-03
41	LepPhi	1.755e-03
42	BJet2Phi	1.558e-03
43	Jet1Phi	1.337e-03

Table 2: Method unspecific separation ranking obtained using BDT method in TMVA

S.N	Var	Sep
1	**loglratio**	4.051e-02
2	LepJetDRmin	3.332e-02
3	LepBJetDRmin	3.285e-02
4	BBDR	2.927e-02
5	**logldihiggs**	2.899e-02
6	LepNeuDPhi	2.694e-02
7	LepPhi	2.636e-02
8	LepJetDRmax	2.611e-02
9	METBJetDPhi_min	2.608e-02
10	Jet1Phi	2.601e-02
11	LepBJetDRmax	2.573e-02
12	BBDPhi	2.565e-02
13	Jet3Phi	2.519e-02
14	**loglttbar**	2.505e-02
15	BJet1Eta	2.502e-02
16	Jet3Eta	2.492e-02
17	Jet2Phi	2.490e-02
18	BJet2Eta	2.482e-02
19	METPhi	2.471e-02
20	METBJetDPhi_max	2.467e-02
21	BJet1Phi	2.460e-02
22	LepEta	2.411e-02
23	BJet2Phi	2.407e-02
24	BJet1Pt	2.399e-02
25	Jet2Eta	2.364e-02
26	**ht2_r**	2.342e-02
27	LepPt	2.314e-02
28	Jet1Eta	2.283e-02
29	**mt2**	2.113e-02
30	**new**	2.047e-02
31	BBM	1.994e-02
32	BJet2Pt	1.954e-02
33	**ht2**	1.935e-02
34	Topness	1.901e-02
35	MET	1.836e-02
36	Jet3Pt	1.722e-02
37	WWMT	1.720e-02
38	Higgsness	1.719e-02
39	Jet1Pt	1.702e-02
40	LepNeuMT	1.614e-02
41	NJets	1.548e-02
42	Jet2Pt	1.345e-02
43	LepM	1.160e-02

Table 3: Method unspecific separation ranking obtained using BDT method in TMVA

References

[1] The ATLAS Collaboration, Phys. Lett. B716 (2012) 1 and ATLAS-CONF-2012-162;The CMS Collaboration, Phys. Lett. B716 (2012) 30 and CMS-PAS-HIG-12-045.

[2] P. Higgs, Phys. Lett. 12 (1964) 132; Phys. Rev. Lett. 13 (1964) 508; F. Englert andR. Brout, Phys. Rev. Lett. 13 (1964) 321; G. Guralnik, C. Hagen and T. Kibble,Phys. Rev. Lett. 13 (1964) 585; S. Weinberg, Phys. Rev. Lett. 19 (1967) 1264.

[3] Baglio, J. et al. "The Measurement of the Higgs Self-Coupling at the LHC: Theoretical Status." Journal of High Energy Physics 2013.4 (2013): n. pag. Crossref. Web.

[4] O. Eboli, G. Marques, S. Novaes and A. Natale, Phys. Lett. B197(1987) 269.

[5] E. Glover and J. van der Bij, Nucl. Phys. B309 (1988) 282.

[6] D. Dicus, C. Kao and S. Willenbrock, Phys. Lett. B203 (1988) 457.

[7] T. Plehn, M. Spira and P.M. Zerwas, Nucl. Phys. B479 (1996) 46;Err.ibidB531(1998) 655.

[8] O. Eboli, G. Marques, S. Novaes and A. Natale, Phys. Lett. B197(1987) 269.

[9] W. Y. Keung, Mod. Phys. Lett. A2 (1987) 765; D. Dicus, K. Kallianpur and S. Willenbrock, Phys. Lett. B200 (1988) 187; K. Kallianpur, Phys. Lett. B215 (1988) 392;A. Dobrovolskaya and V. Novikov, Z. Phys. C52 (1991) 427.

[10] A. Abbasabadi, W. Repko, D. Dicus and R. Vega, Phys. Rev. D38(1988) 2770.

[11] V. Barger, T. Han and R. Phillips, Phys. Rev. D38 (1988) 2766.[15] M. Moretti et al., JHEP 0502 (2005) 024

[12] M. Moretti et al., JHEP 0502 (2005) 024.

[13] *Detector & Technology* , available at `https://atlas.cern/discover/detector`.

[14] "Overall detector concept". ATLAS Technical Proposal. CERN. 1994.

[15] *The Inner Detector* , available at `https://atlas.cern/discover/detector/inner-detector`.

[16] *Monte Carlo method*, available at `https://en.wikipedia.org/wiki/Monte`$_{C}arlo_{m}ethod$.

[17] ATLAS Collaboration, The ATLAS Simulation Infrastructure, Eur. Phys. J. C 70 (2010) 823, arXiv: 1005.4568 [physics.ins-det].

[18] *Calorimeters*, available at `https://atlas.cern/discover/detector/calorimeter`.

[19] Aad, G. et al. "Search for Non-Resonant Higgs Boson Pair Production in the Bb Final State with the ATLAS Detector in Pp Collisions at S=13 TeV." Physics Letters B 801 (2020): 135145. Crossref. Web.

[20] *Decision Tree*, available at https://en.wikipedia.org/wiki/Decision_tree

[21] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, TMVA - Toolkit for Multivariate Data Analysis, PoS ACAT 040 (2007), arXiv:physics/0703039

[22] *Curse of dimensionality* , available at https://en.wikipedia.org/wiki/Curse_of_dimensionality.

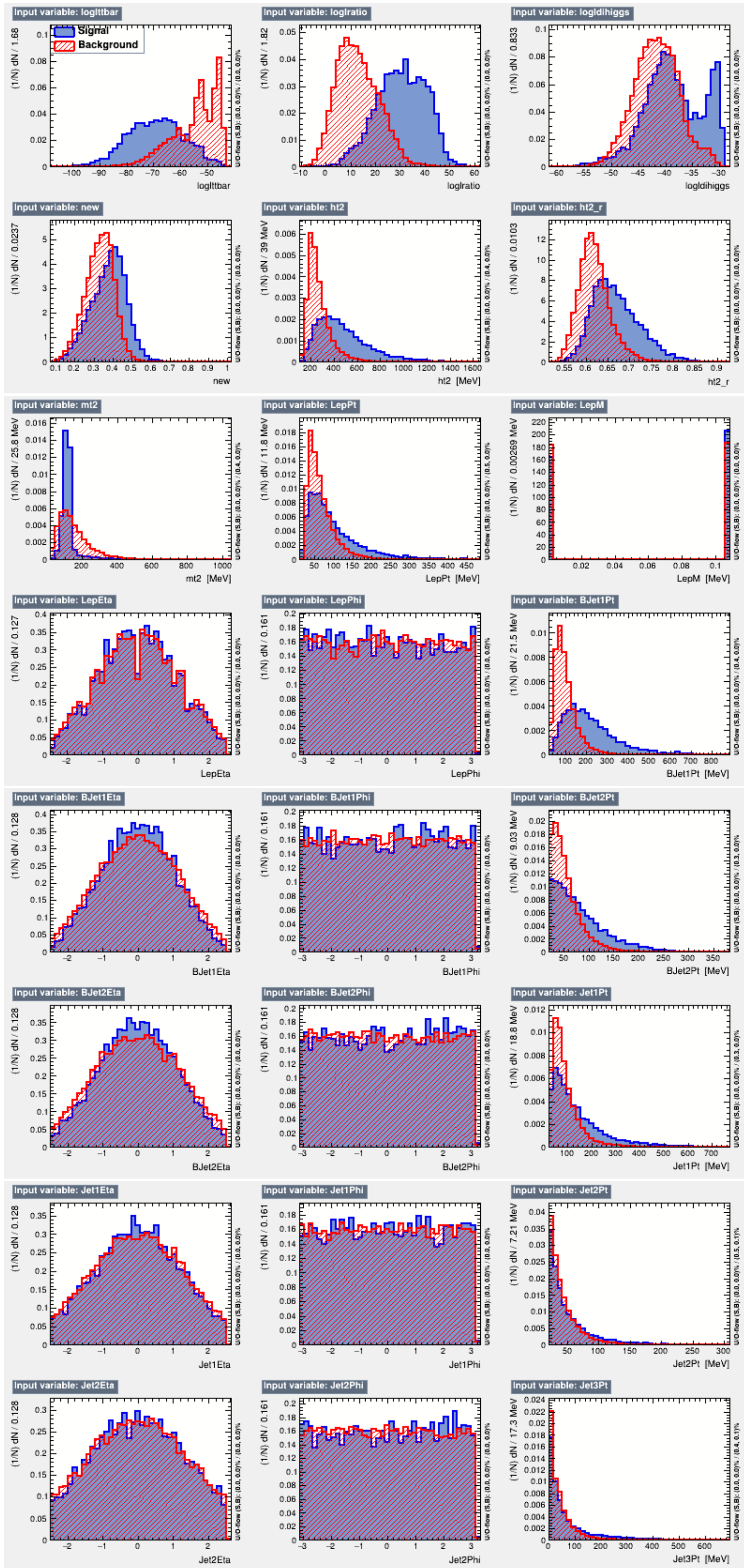
[23] *Dimensionality reduction* , available at https://en.wikipedia.org/wiki/Dimensionality_reduction.

[24] Sitong An, University of Cambridge, U.K., DESY ATLAS Group, "Optimising Variable Selection for Machine Learning Analysis in ATLAS ttH Search", September 7, 2017. Available at https://www.desy.de/f/students/2017/reports/SitongAn.pdf.gz.

[25] Sjöstrand, Torbjörn et al. "An Introduction to PYTHIA 8.2." Computer Physics Communications 191 (2015): 159–177. Crossref. Web.

[26] *The POWHEG BOX user manual: common features*, available at https://usermanual.wiki/Pdf/manualBOX.1264151446.pdf.

A Appendix


Input variable: logttbar
Signal
Background
Input variable: logratio
Input variable: logdihiggs
Input variable: new
Input variable: ht2
Input variable: ht2_r
Input variable: mt2
Input variable: LepPt
Input variable: LepM
Input variable: LepEta
Input variable: LepPhi
Input variable: BJet1Pt
Input variable: BJet1Eta
Input variable: BJet1Phi
Input variable: BJet2Pt
Input variable: BJet2Eta
Input variable: BJet2Phi
Input variable: Jet1Pt
Input variable: Jet1Eta
Input variable: Jet1Phi
Input variable: Jet2Pt
Input variable: Jet2Eta
Input variable: Jet2Phi
Input variable: Jet3Pt

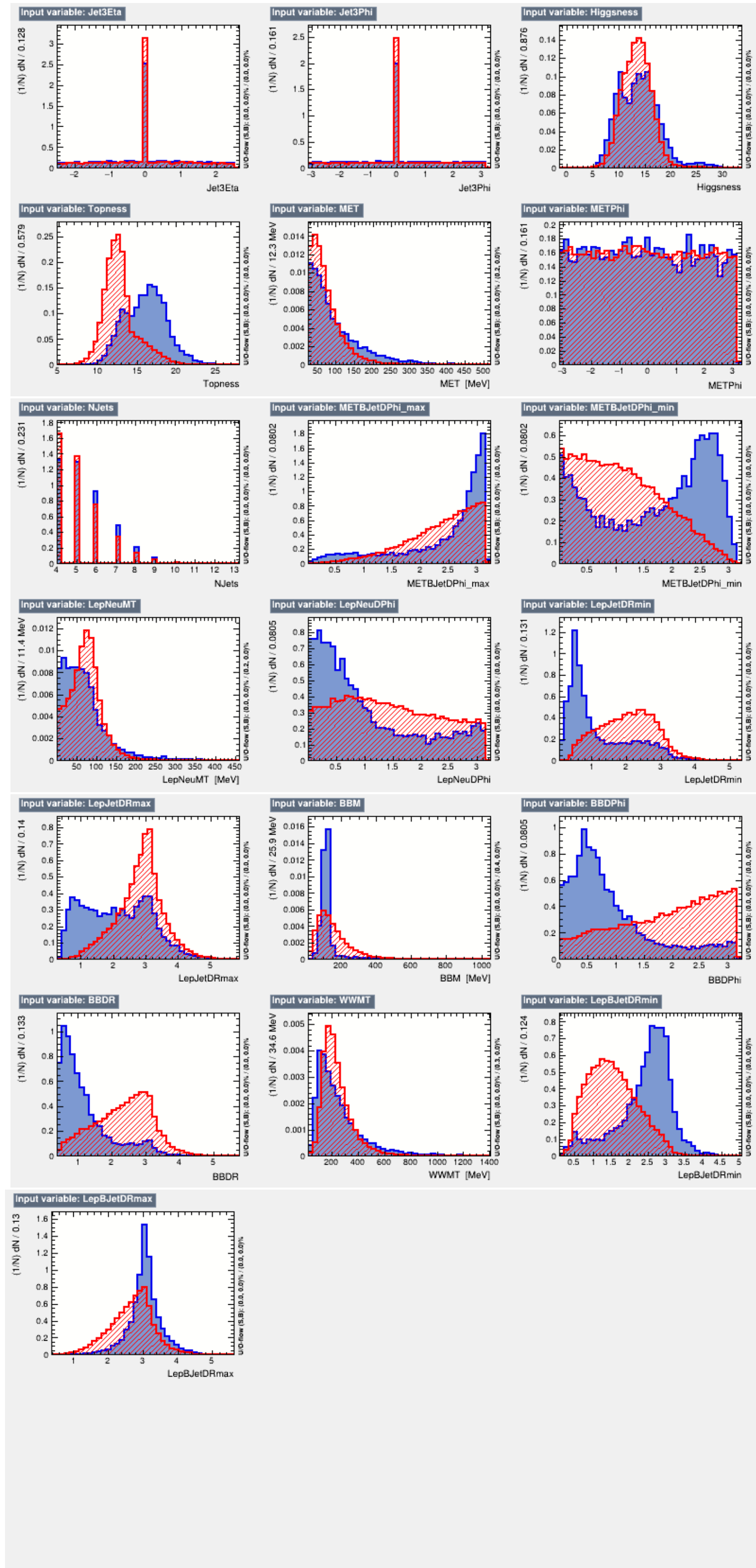


Figure 15: Distribution of 43 trained variables

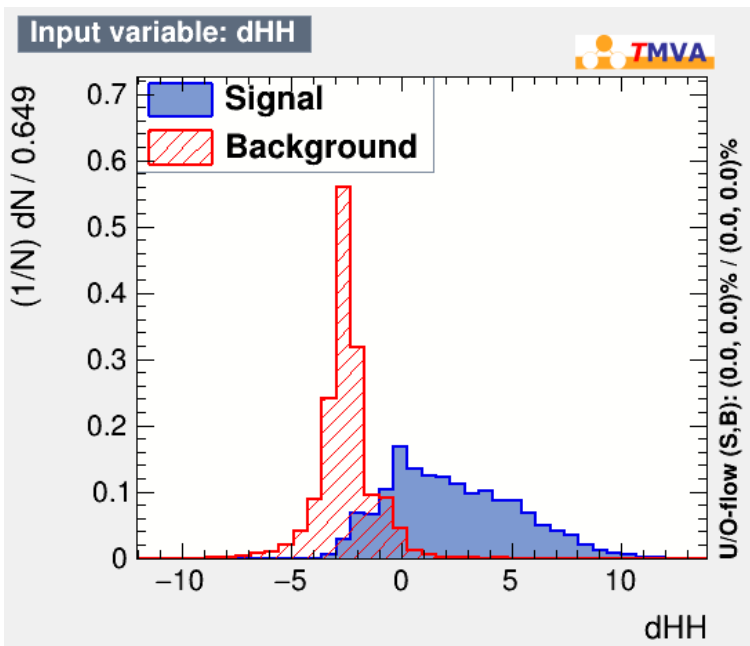


Figure 16: Distribution of the output of 36 DNN trained variables, dHH.