Application of deep neural networks in Top-Higgs coupling analysis with the ATLAS detector
Применение глубоких нейронных сетей для анализа топ-Хиггс константы связи с детектором ATLAS

A. Didenko[a1], *I. Yeletskikh*[a]

А.Р. Диденко[a,1], *И.В. Елецких*[a]

[a] Joint Institute for Nuclear Research, Dubna, Moscow Region, Russia, 141980

Рождение бозона Хиггса в ассоциации с одним t-кварком позволяет напрямую измерить относительную фазу и величину константы связи Хиггса с t-кварком. Из-за высокой схожести маркеров идентификации конечного состояния с основными фоновыми процессами и чрезвычайно малого числа сигнальных событий для достижения достаточного разделения сигнала и фона требуется применение методов машинного обучения. В этой работе представлены результаты применения метода глубоких нейронных сетей (DNN), который использовался в качестве метода перекрёстной проверки для классификации событий в анализе АТЛАС.

The associated production of a Higgs boson with a single top quark provides direct sensitivity to the relative phase and magnitude of the Higgs–top coupling. Due to the high similarity of the final–state signature with the dominant backgrounds and extremely small number of signal events, machine learning techniques are required to achieve sufficient signal–background separation. Results of Deep Neural Network (DNN) approaches as a cross-check method for event classification problem in the ATLAS analysis are presented in this paper.

PACS: 07.05Mh; 14.80.Bn

Introduction

The top quark possesses the strongest coupling strength to the Higgs boson in Standart Model (SM). The single-top-Higgs production process (tH) provides direct sensitivity to the relative phase of the complex tH Yukawa coupling, making it a target for testing the SM predictions [1–3]. The cross section of tH with one light-flavor quark production channel (tHq) in SM is 74.3 fb^{-1}, which is 7 times smaller than $t\bar{t}H$ cross section [1,4]. In addition to that, the kinematics of signal and background processes are very close. The separation power of individual kinematic variables is insufficient to classify the rare Higgs-top process. The markers of tHq signal with Higgs decay into $b\bar{b}$ include more than 2 b-jets with an extra soft one, a light jet from light-flavor quark (dominantly a u-quark), 1 tight lepton and missing transverse momentum from W-boson decay (see Figure 1) [5].

[1]E-mail: alisadidenko@jinr.ru

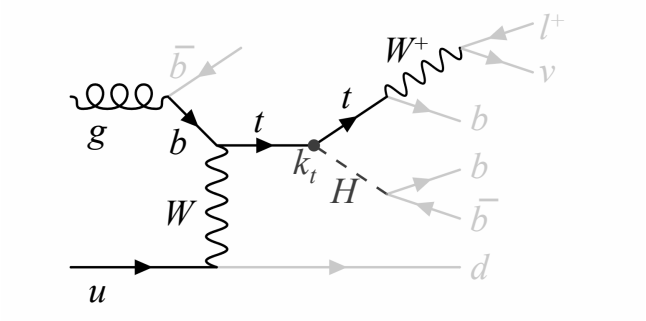


Fig. 1: Dominant Feynman diagram for tHq production channel with Higgs decay to b-quarks. The final–state signature of $tHq(H \to b\bar{b})$ is presented by light-gray, and the Higgs–top Yukawa coupling is shown by dark-gray.

Deep neural network method

Deep Neural Networks (DNN), were developed and tested for ATLAS analysis with Run 2 data [5] as a cross-checks of the main analysis method based on Boosted Decision Trees (BDT). The events are selected using the same pre-selection criteria as for BDT analysis. After pre-selection, simulated and experimental event yields are normalized to the integrated luminosity. Only the tHq, $t\bar{t}+ \geq 1b$ (ttb), $t\bar{t}+ \geq 1c$ (ttc), $t\bar{t}+ \geq 0$ $light$ (ttL) and $t\bar{t}H$ (ttH) production processes were used for DNN learning. The most sensitive kinematic variables are used for classification tHq signal and three main background components ttb, ttc and ttL. Highly correlated kinematic features were excluded (more than 84%) to prevent redundancy.

The Keras [6] API within the TMVA [7] framework were used for built and trained a DNN models. A fully-connected feedforward multi-layer perceptron (MLP) architectures were used for multiclass classification separating signal and the three major backgrounds. Input layer contains 24 neurons for tHq, 24 neurons for ttb, 28 neurons for ttc, and 23 neurons for ttL, which are correspond to the input kinematic variables for each of binary classifiers. Architecture of two hidden layers are the following: 150 and 150 neurons for tHq, 240 and 240 neurons for ttb, 150 and 150 neurons for ttc, and 230 neurons for the first hidden layer and only 23 neurons for the second one in ttL-classifier. It was trained in a supervised manner on Monte Carlo samples with Adam optimizer using 20 number of learning iterations. Hyperparameters were optimized manually. The output neurons as a final DNN responses are presented in Figure 2.

Results

The applying of DNN allows to increase the significance of the SM signal by 1.4 times (see Figure 3). The maximum significance of tHq is achieved with cut on tH-neuron response > 0.82. The separation power of tH-neuron equals to 22.42%.

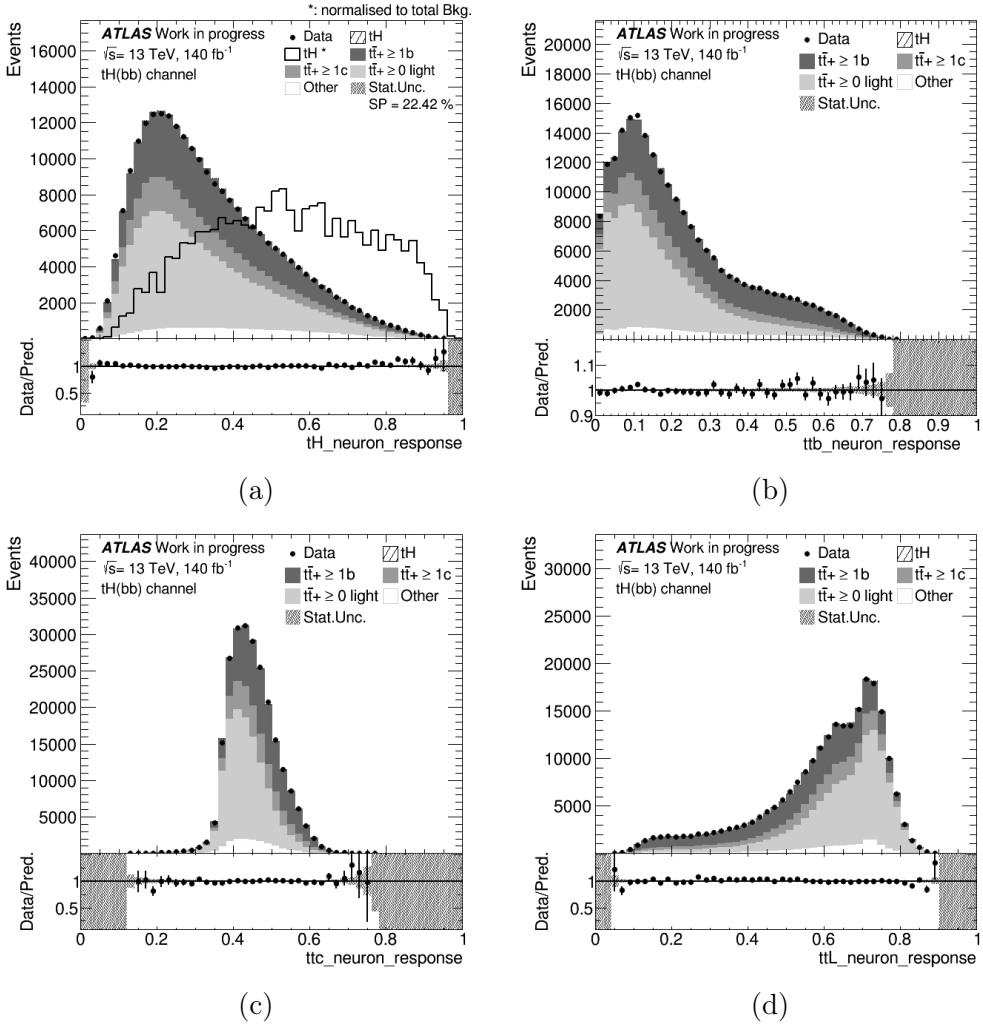


Fig. 2: Neural network response to data events and simulated data for signal neuron (a), *ttb* (b), *ttc* (c), and *ttL* (d). Bad separation for *ttc* is due to its kinematic properties, not to DNN learning. The other background process include ttH, tZq, ttZ, ttW, tWZ, W,$Z + jets$, tW, single t and non-top background processes.

The DNN multiclassifier are used to separate different background components to better constrain them, as well as to define the analysis regions. The result of the fit in the control regions determined by the multiclassifier: 1.25 for *ttb*, 1.1 for *ttc*, and 0.98 ttL.

Comparison of DNN and BDT signal-score distributions are provide using signal (SR) and control (CR) regions (see Figure 4), which are defined as orthogonal kinematic regions with the highest significance tHq signal and *ttb* contributions, respectively [5]. The 34 number of tHq events are expected in SR by DNN prediction (see Table 1) instead of 43 ± 7 by prediction of BDT classifier for pre-fit yields [5].

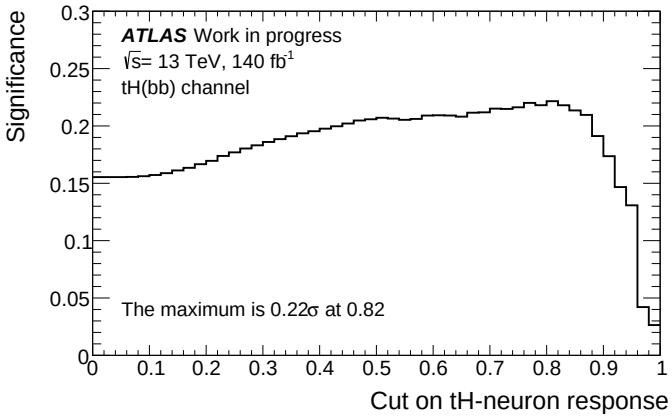
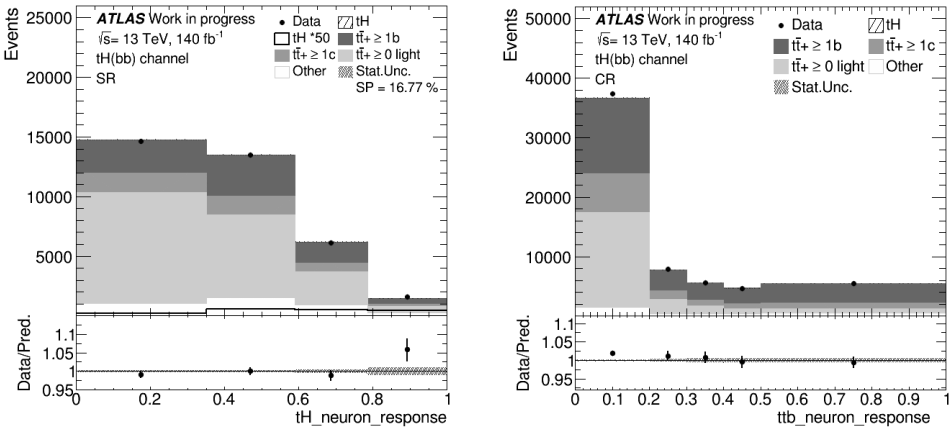


Fig. 3: The signal significance for the tH process.



(a) (b)

Fig. 4: (a) DNN tH-neuron response in signal region; (b) DNN ttb-neuron response in control region.

Conclusion

Deep machine learning shows good performance w.r.t. signal/background separation in the analysis of the $tHq(H \to b\bar{b})$ channel at the ATLAS experiment. The DNN separation efficiency comparable to the basic analysis method BDT. After application of the DNN model the signal significance increases by a factor of 1.4. Future developments of algorithms for optimization of DNN hyperparameters can improve sensitivity to the tH process.

REFERENCES

1. *F. Demartin, B. Maier, F. Maltoni et al.* tWH associated production at the LHC // Eur. Phys. J. C.— 2017.— V. 77 (34).— https://doi.org/10.1140/epjc/s10052-017-4601-7.

Process	Pre-selection yield before cut on tH-neuron	After cut on tH-neuron
tH	74	9
$(t\bar{t}+ \geq 1b) \times 1.24$	74447	688
$t\bar{t}+ \geq 1c$	33516	194
$t\bar{t}+ \geq 0$ *light*	106055	509
Others	16231	261
Total	230326	1661
Data	236832	1867

Table 1: Expected event yields after the pre-selection for the signal and background processes at the LHC, for an integrated luminosity of 140 fb^{-1} at $\sqrt{s} = 13$ TeV.

2. *ATLAS Collaboration.* Probing the CP nature of the top-Higgs Yukawa coupling in $t\bar{t}H$ and tH events with $H \to b\bar{b}$ decays using the ATLAS detector at the LHC // Phys. Lett. B. — 2024. — V. 849. — P. 138469. — https://doi.org/10.1016/j.physletb.2024.138469.

3. *F. Demartin, B. Maier, F. Maltoni et al..* Measurement of the properties of Higgs boson production at $\sqrt{s} = 13$ TeV in the $H \to \gamma\gamma$ channel using 139 fb^{-1} of pp collision data with the ATLAS experiment // J. High Energ. Phys. — 2023. — V. 7, no. 88. — https://doi.org/10.1007/JHEP07(2023)088.

4. *Grojean C.* Higgs Physics // CERN Yellow Reports. — 2016. — V. 5. — https://doi.org/10.5170/CERN-2016-005.143.

5. *ATLAS Collaboration.* Search for the production of a Higgs boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector // J. High Energ. Phys. — 2025. — V. 10, no. 93. — https://doi.org/10.1007/JHEP10(2025)093.

6. *Chollet F.* Keras [Electronic resource]. — https://keras.io/– Access mode: free. — 2015.

7. *et al. A.H.* Toolkit for Multivariate Data Analysis [Electronic resource]. — https://root.cern/manual/tmva/– Access mode: free. — 2007.