

Tagging large-radius b -jets from Higgs decays dropping unneeded information

Andrea Di Luca,^{a,b,c,*} Daniela Mascione,^{a,b,c} Francesco Maria Follega,^{a,b} Marco Cristoforetti^{b,c} and Roberto Iuppa^{a,b}

^a*Dipartimento di Fisica, Università di Trento, Via Sommarive 14, 38123 Trento, Italy*

^b*TIFPA, Via Sommarive 14, 38123 Trento, Italy*

^c*FBK, Via Sommarive 18, 38123 Trento, Italy*

E-mail: andrea.diluca@unitn.it

Multivariate approaches used in physics analyses by the High Energy Physics community often combine high-level observables estimated by very complex algorithms. The process to select these variables is usually based on a “brute force” approach, where all available event features are tested for multiple combinations of the algorithm hyperparameters.

In this work, we propose an original method based on the use of a CancelOut layer to select to give as input to a Fully Connected Neural Network. Promising results are obtained in the development of a DNN classifier to select proton-proton collisions where a boosted Higgs boson decay to two b -quarks.

*** *Particles and Nuclei International Conference - PANIC2021* ***

*** *5 - 10 September, 2021* ***

*** *Online* ***

*Speaker

1. Introduction

In the high-energy physics community, there is a growing interest in replacing cut-based selections using different types of multivariate analysis. This transformation opened up to use high-level variables produced by complex reconstruction algorithms. Within this context, Deep Learning approaches are rapidly spreading, combining all the available information and improving the selection performances. The development of these algorithms often relies on a trial and error approach, where all available event features are tested for multiple combinations of the algorithm hyperparameters. Herein, we propose an effective method to choose the most valuable variables to give as input to a Deep Neural Network using a CancelOut layer. Indeed, given a fixed number of variables, the CancelOut layer selects during training only the most relevant features to achieve the best performances.

2. CancelOut layer

The CancelOut layer [1] is composed of neurons that have one single input. The node's output is the product of the input feature and the sigmoid of the associated weight. After training, irrelevant features will have an associated CancelOut weight that outputs small numbers after the sigmoid application. In contrast, for the relevant features, the output will be close to one. In its original formulation, the weights of the CancelOut layer are used as a metric to evaluate the impact of one feature in the decision process of a DNN model. Our purpose is different: we developed a modified architecture of the CancelOut layer so that only a certain fraction, defined by the user, of the CancelOut nodes are active (sigmoid of the weight equal to 1) while the others are switched off (sigmoid of the weight equal to 0).

3. Boosted $H \rightarrow bb$ tagging

We tested the CancelOut layer while developing a fully connected DNN to classify pp collision events where a Higgs boson with very high transverse momentum decays to two b -quarks. In this regime, the decay products of the Higgs boson are very collimated and it is challenging to resolve the di-jet structure [2]. A single large and massive jet containing both the b -quarks originated jets are more likely to be reconstructed. The $H \rightarrow bb$ decay channel is interesting for the study of Higgs boson properties since it accounts for 58% of the total Higgs boson decays [3]. Moreover, larger deviations from Standard Model predictions are expected in the boosted regime [4]. However, recognizing these events in a pp collision experiment represents a challenging task, mainly because of the huge irreducible background of QCD multi-jet production.

3.1 Simulated data and object reconstruction

The dataset used to develop the classifier is produced using a framework developed by combining Pythia8[5], to generate high-energy physics events, Delphes [6], to simulate the detector response and RAVE [7] for secondary vertex reconstruction. Large radius anti- k_t jets [8] (large- R jets) with $R = 1$ are reconstructed together with variable radius track jets [9] with $R_{\text{MAX}} = 0.4$, $R_{\text{MIN}} = 0.02$ and $\rho = 30$ GeV (ρ parameter determines how fast the effective jet size decreases

with the transverse momentum of the jet). For the large-R jets, we defined kinematic variables plus jet substructure variables. For the variable R track jets, we defined kinematic variables plus the b-tagging information and variables connected to the secondary vertex. We selected large-R jets with $p_T > 450 \text{ GeV}/c^2$ and $\eta < 2$. Then we look for the 2 highest p_T track jets contained in a selected large-R jet. The total number of initial features is 39.

3.2 Reduction of the number of input features

Different models were trained by varying the number of desired features that the CancelOut layer must activate. The top plot in fig 1 shows the behavior of the CancelOut weight associated with each feature when varying the number of requested features. Relevant features are activated ($\sigma(w) \simeq 1$) already when asking for a low number of features. Irrelevant features, which have not been activated, do not take part in the calculation since they have $\sigma(w) \simeq 0$. Then, increasing the number of asked features one by one, single features are activated and remain active in most cases. The performances of the models are shown on the bottom side of fig. 1. It is important to notice that after a certain number of features are switched on, there is no significant improvement in the performance. This confirms the idea that irrelevant features were last activated.

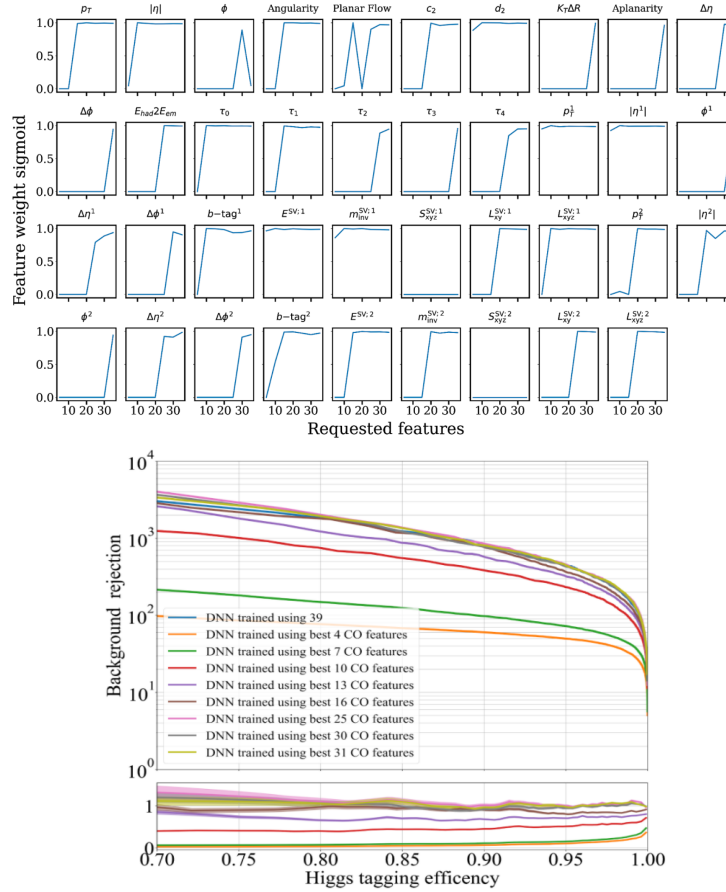


Figure 1: (Top) CancelOut weight for each feature as function of the desired number of features. (Bottom) Background rejection rate versus Higgs tagging efficiency with large-R jets of different model trained by varying the number of features.

4. Conclusion

In this work, we proposed an original method based on the use of the CancelOut layer to select features in input to the Neural Network. We used as a study case the selection of events where a boosted large and massive jet contains both of the b quarks originated from H boson decay. We showed how the Deep Neural Network classifier performance can be affected by keeping irrelevant variables in input and how our method can naturally get rid of them. The proposed method can be easily implemented in already developed Deep Neural Network classifiers through a retraining campaign.

References

- [1] V. Borisov, J. Haug and G. Kasneci, *Cancelout: A layer for feature selection in deep neural networks*, in *Artificial Neural Networks and Machine Learning – ICANN 2019: Deep Learning*, I.V. Tetko, V. Kůrková, P. Karpov and F. Theis, eds., (Cham), pp. 72–83, Springer International Publishing, 2019.
- [2] ATLAS COLLABORATION collaboration, *Performance of large- R jets and jet substructure reconstruction with the ATLAS detector*, Tech. Rep. [ATLAS-CONF-2012-065](#), CERN, Geneva (Jul, 2012).
- [3] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, [1610.07922](#).
- [4] M. Grazzini, A. Ilnicka, M. Spira and M. Wiesemann, *Modeling BSM effects on the Higgs transverse-momentum spectrum in an EFT approach*, *JHEP* **03** (2017) 115 [[1612.00283](#)].
- [5] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [[1410.3012](#)].
- [6] DELPHES 3 collaboration, *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057 [[1307.6346](#)].
- [7] W. Waltenberger and F. Moser, *Rave - an open, extensible, detector-independent toolkit for reconstruction of interaction vertices*, in *2006 IEEE Nuclear Science Symposium Conference Record*, vol. 1, pp. 104–109, 2006.
- [8] M. Cacciari, G.P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *Journal of High Energy Physics* **2008** (2008) 063.
- [9] D. Krohn, J. Thaler and L.-T. Wang, *Jets with Variable R* , *JHEP* **06** (2009) 059 [[0903.0392](#)].