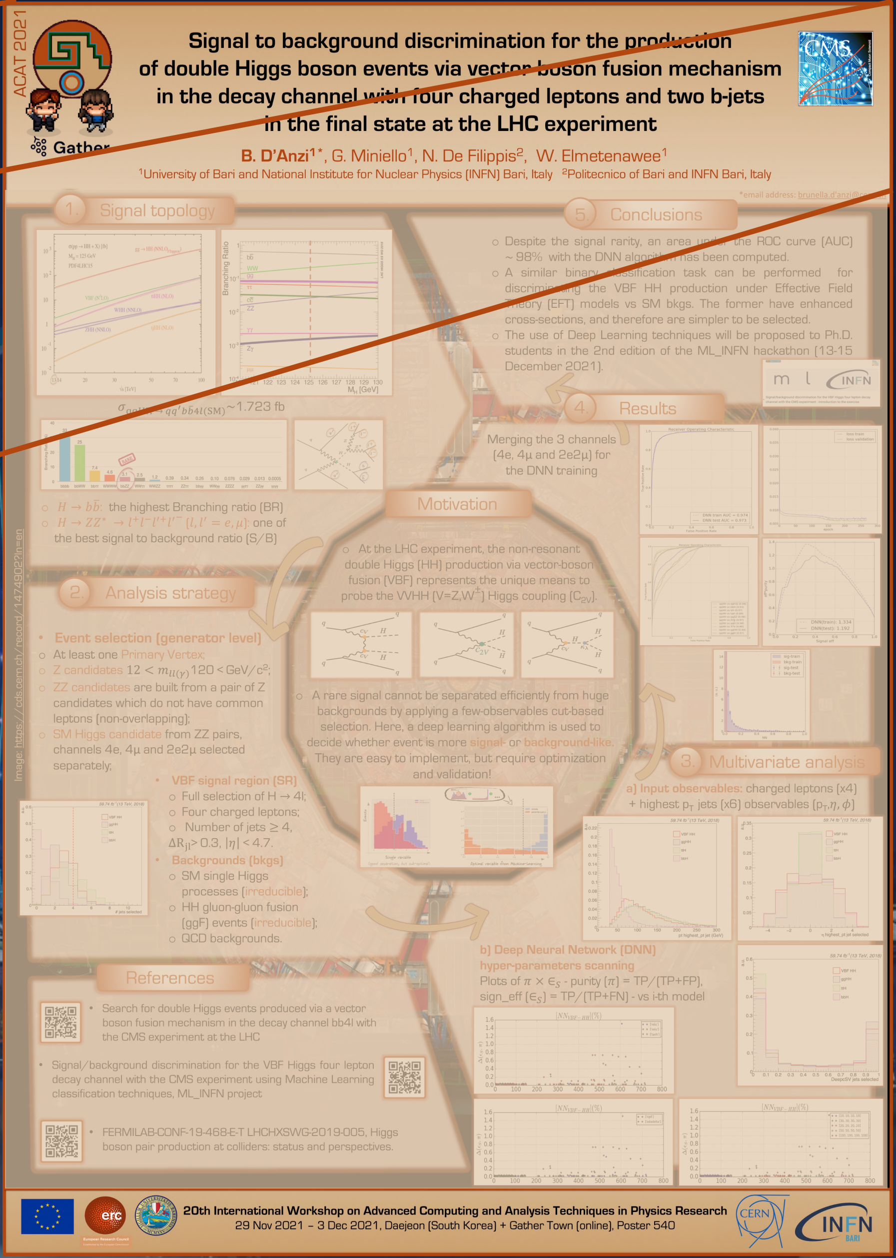
ACAT 2021

Signal to background discrimination for the production of double Higgs boson events via vector boson fusion mechanism in the decay channel with four charged leptons and two b-jets in the final state at the LHC experiment

B. D'Anzi[1]*, G. Miniello[1], N. De Filippis[2], W. Elmetenawee[1]

[1]University of Bari and National Institute for Nuclear Physics (INFN) Bari, Italy [2]Politecnico of Bari and INFN Bari, Italy




B. D'Anzi

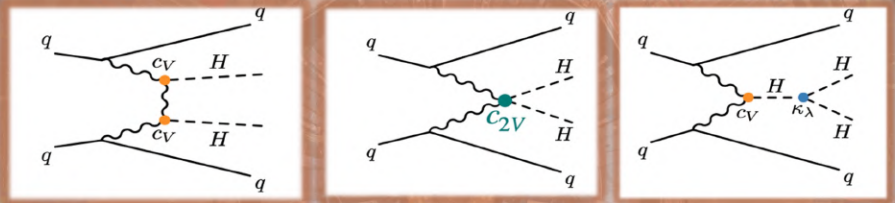

G. Miniello


N. De Filippis

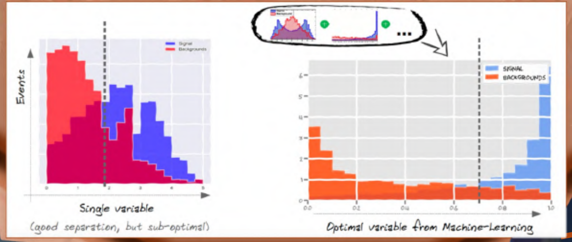

W. Elmetenawee

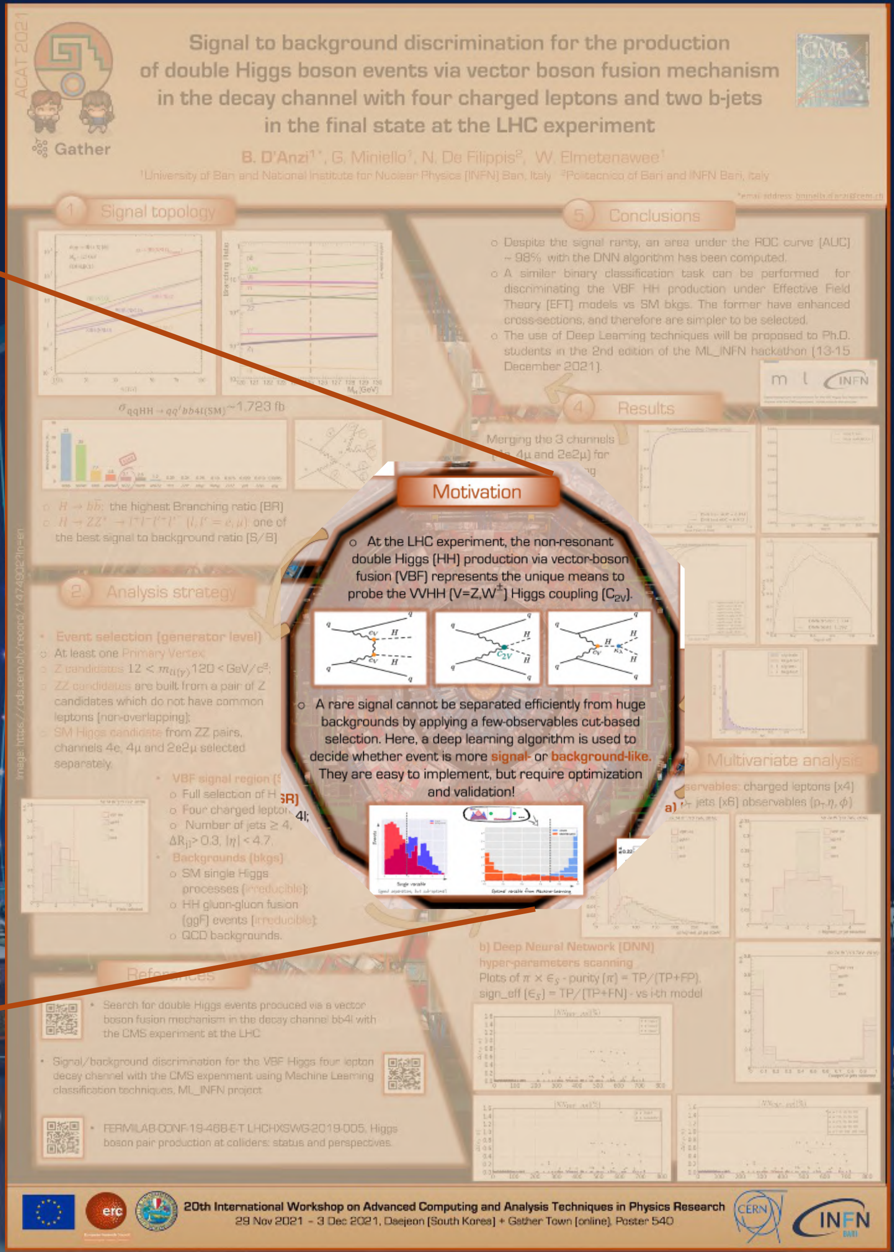
Motivation

- At the LHC experiment, the non-resonant double Higgs (HH) production via vector-boson fusion (VBF) represents the unique means to probe the VVHH ($V=Z,W^{\pm}$) Higgs coupling (C_{2V}).

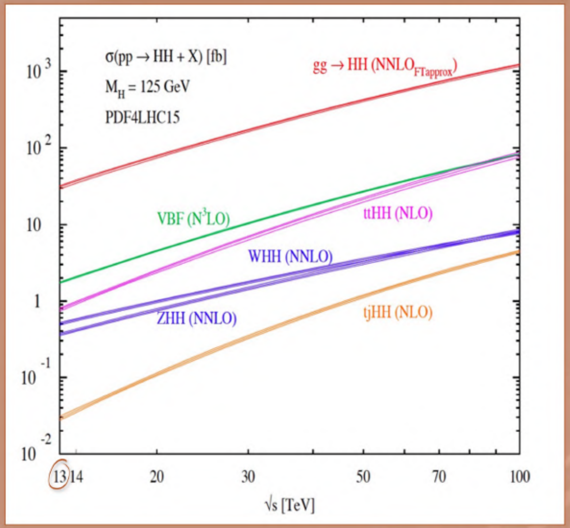


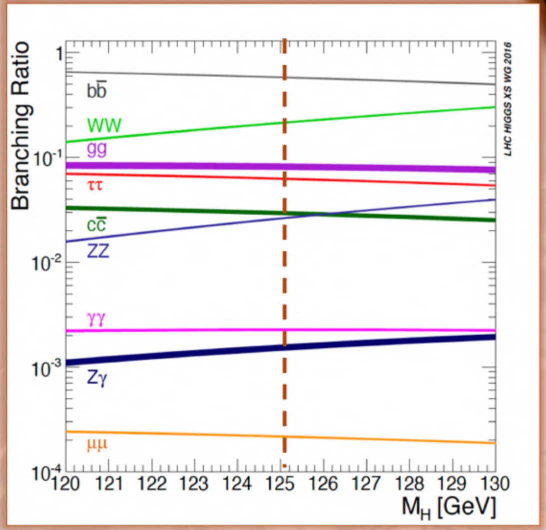
- A rare signal cannot be separated efficiently from huge backgrounds by applying a few-observables cut-based selection. Here, a deep learning algorithm is used to decide whether event is more **signal-** or **background-like.** They are easy to implement, but require optimization and validation!



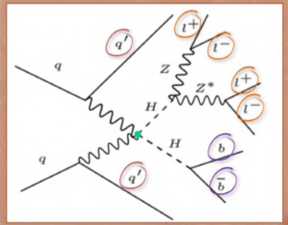


1. Signal topology





$\sigma_{qqHH \rightarrow qq'b\bar{b}4l(SM)}$~1.723 fb



- $H \rightarrow b\bar{b}$: the highest Branching ratio (BR)
- $H \rightarrow ZZ^* \rightarrow l^+l^-l'^+l'^-$ $(l, l' = e, \mu)$: one of the best signal to background ratio (S/B)

At the l

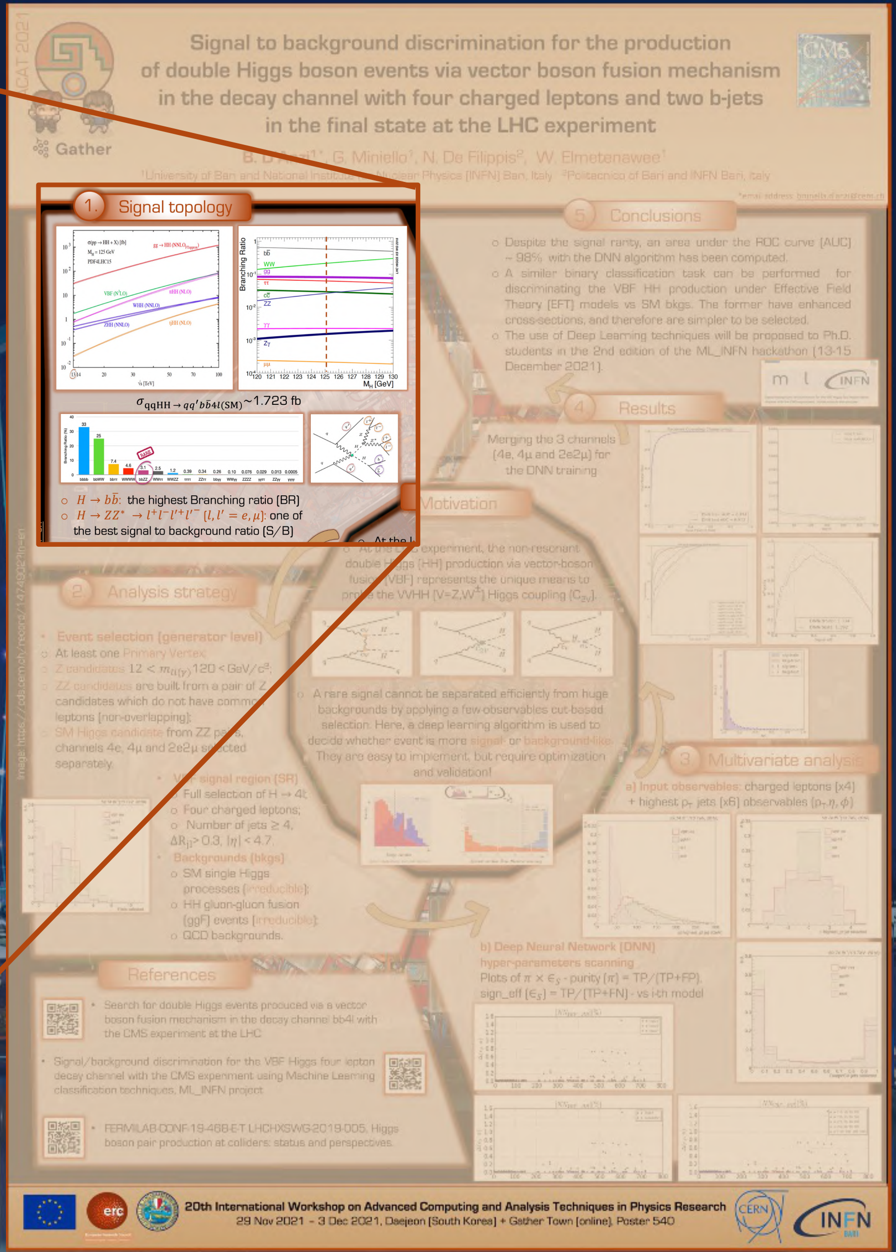
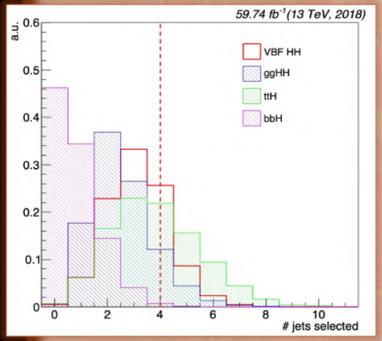


Image: https://cds.cern.ch/record/147490

2. Analysis strategy

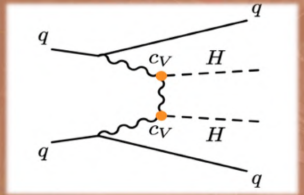
- **Event selection (generator level)**
  - At least one Primary Vertex;
  - Z candidates $12 < m_{ll(\gamma)} 120 < \mathrm{GeV}/c^2$;
  - ZZ candidates are built from a pair of Z candidates which do not have common leptons (non-overlapping);
  - SM Higgs candidate from ZZ pairs, channels 4e, 4μ and 2e2μ selected separately;



- **VBF signal region (SR)**
  - Full selection of $H \rightarrow 4l$;
  - Four charged leptons;
  - Number of jets ≥ 4, $\Delta R_{jl} > 0.3$, $|\eta| < 4.7$.
- **Backgrounds (bkgs)**
  - SM single Higgs processes (irreducible);
  - HH gluon-gluon fusion (ggF) events (irreducible);
  - QCD backgrounds.

fusion (VBF) rep

probe the VVHH



- A rare signal cannot b

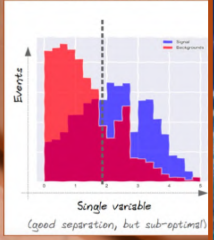
backgrounds by apply

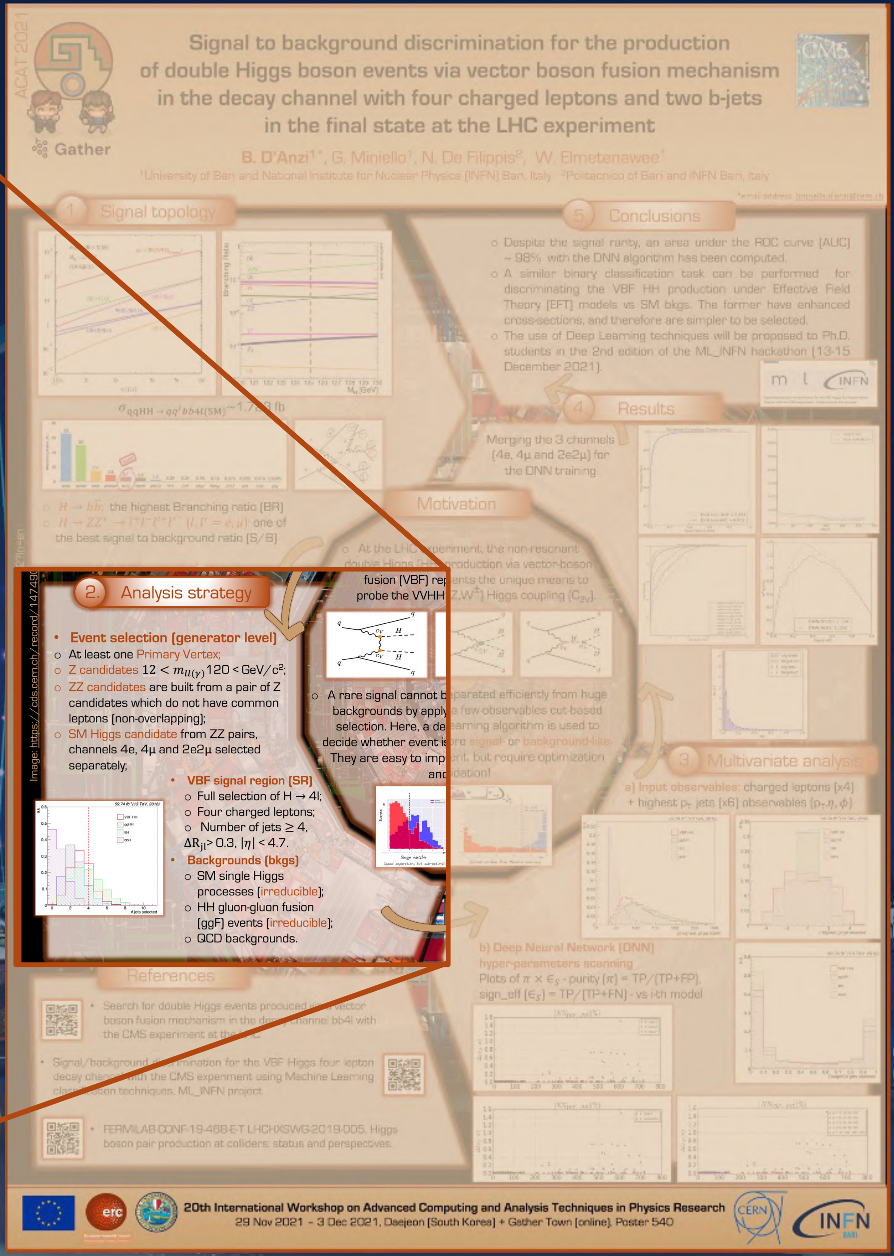
selection. Here, a de

decide whether event is

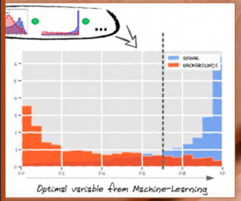
They are easy to imp

and



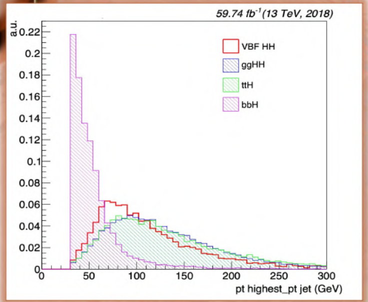


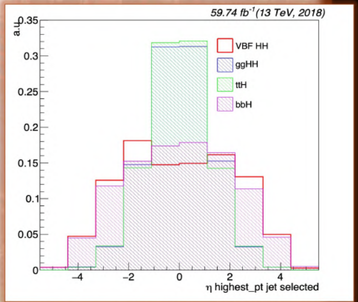
re signal- or background-like.
nt, but require optimization
dation!

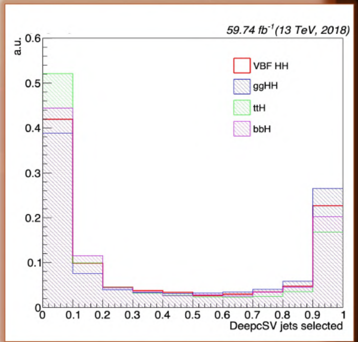


3. Multivariate analysis

a) Input observables: charged leptons (x4)
+ highest p_T jets (x6) observables (p_T,η, ϕ)

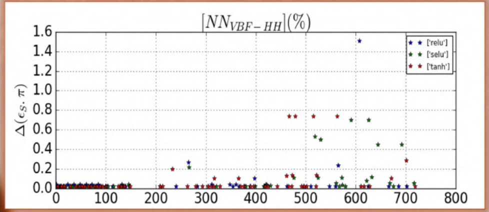


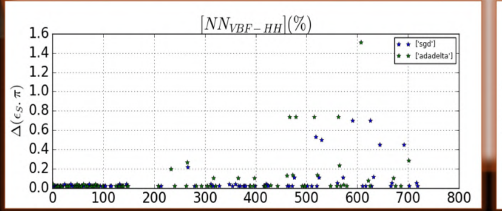


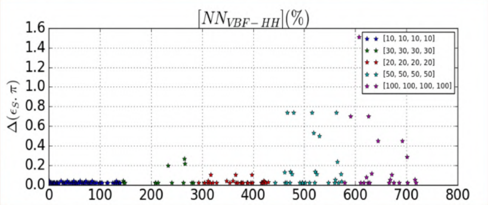


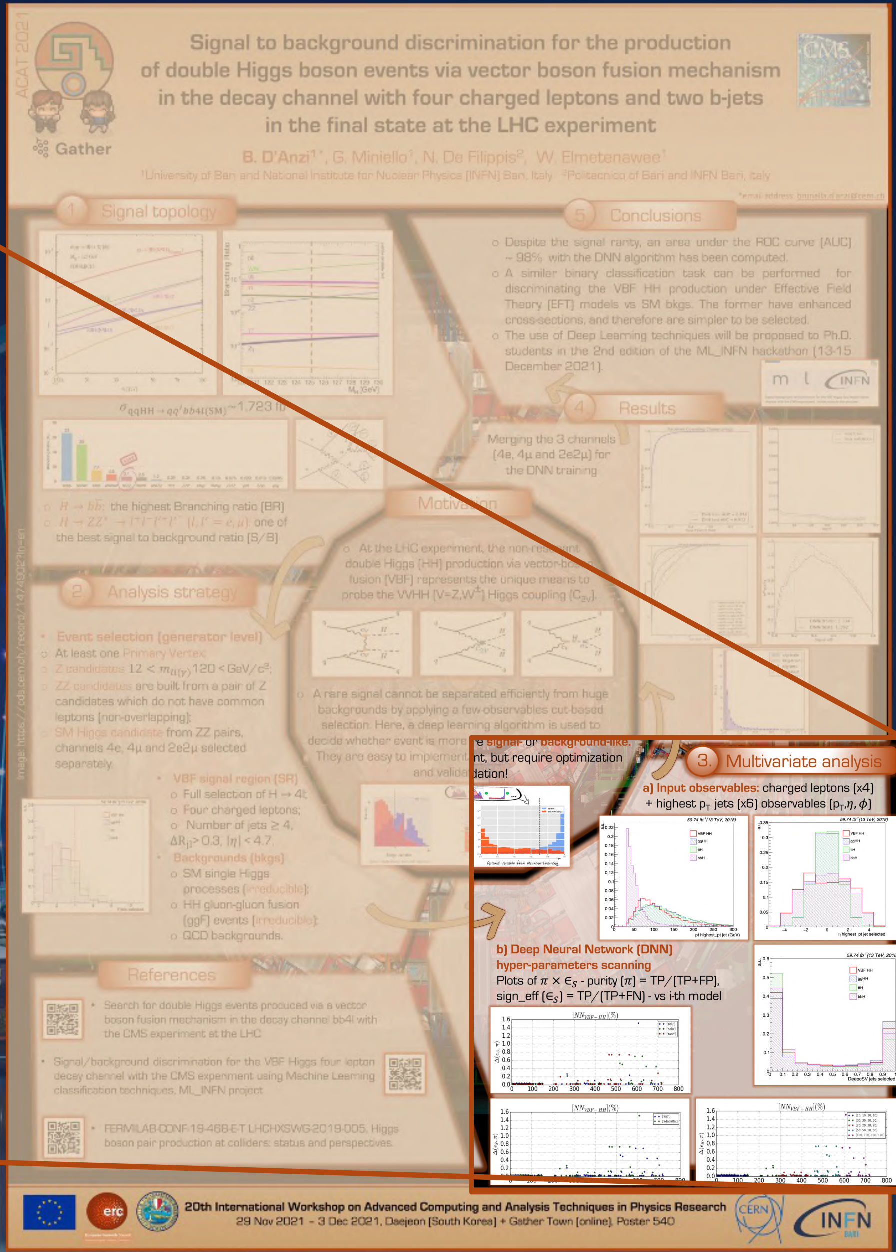
b) Deep Neural Network (DNN) hyper-parameters scanning

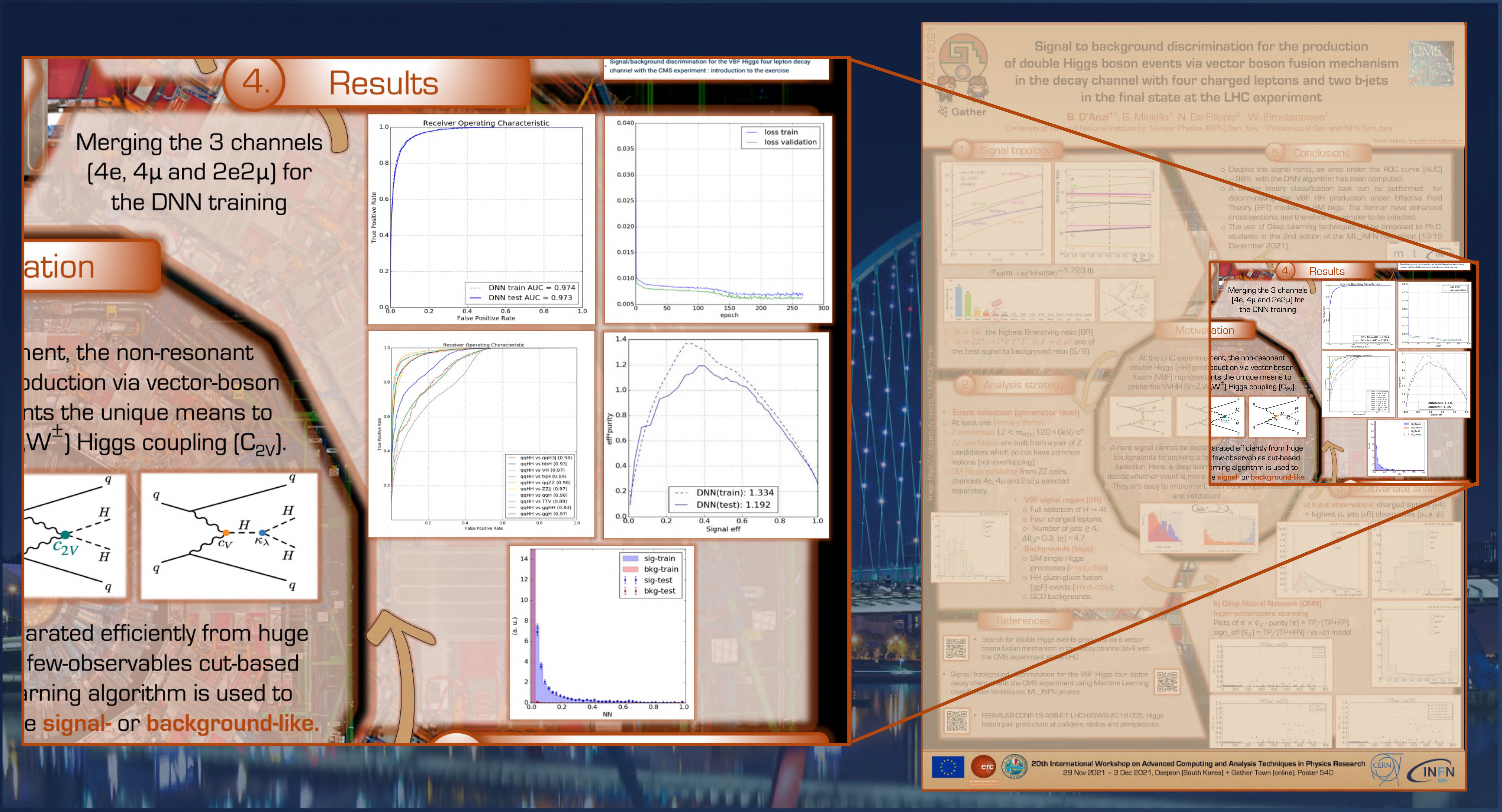
Plots of $\pi \times \epsilon_S$ - purity (π) = TP/(TP+FP), sign_eff (ϵ_S) = TP/(TP+FN) - vs i-th model









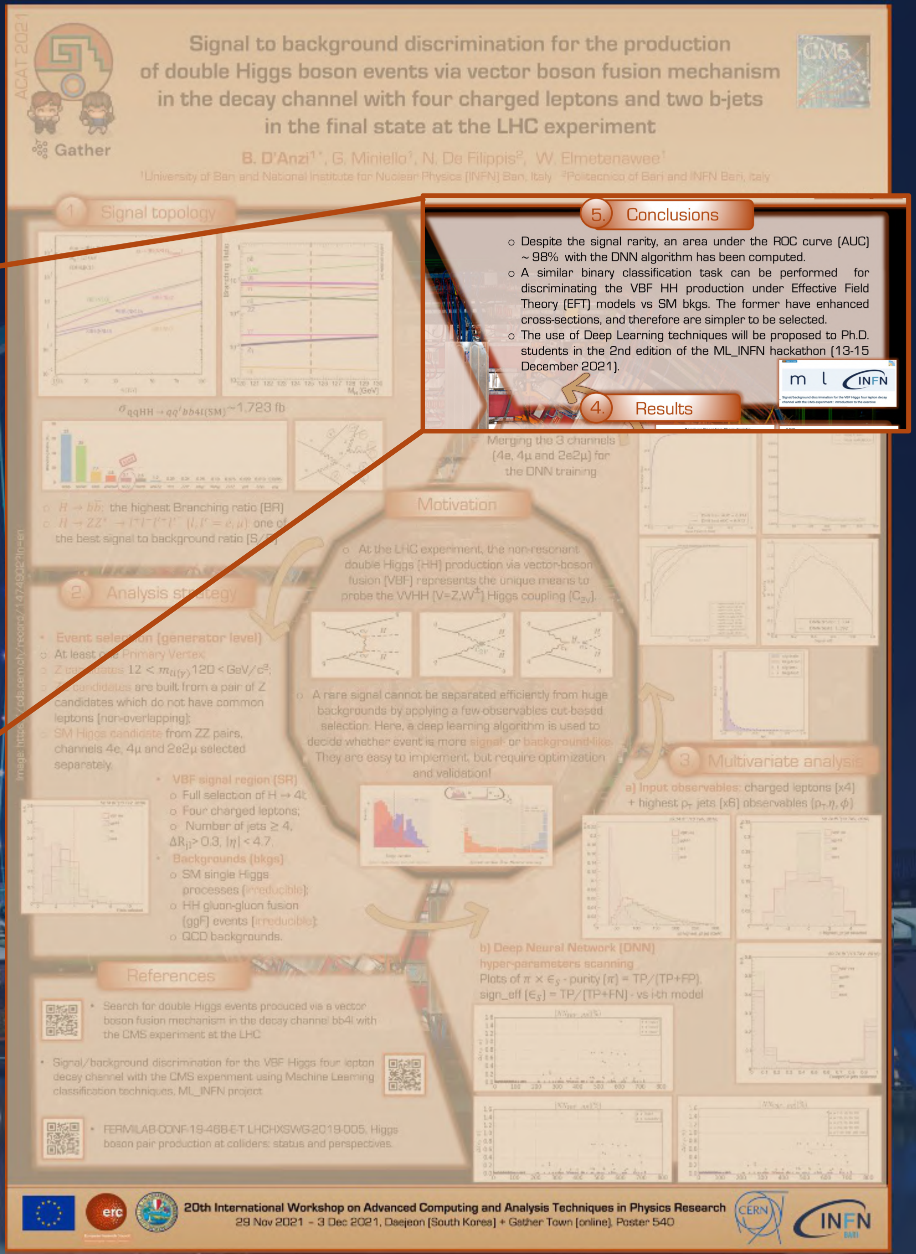

4. Results
Merging the 3 channels (4e, 4μ and 2e2μ) for the DNN training
Signal/background discrimination for the VBF Higgs four lepton decay channel with the CMS experiment : introduction to the exercise
Receiver Operating Characteristic
True Positive Rate
False Positive Rate
DNN train AUC = 0.974
DNN test AUC = 0.973
loss train
loss validation
epoch
qqHH vs qqH3J (0.98)
qqHH vs bbH (0.93)
qqHH vs VH (0.97)
qqHH vs tqH (0.89)
qqHH vs qqZZ (0.98)
qqHH vs ZZjj (0.97)
qqHH vs qqH (0.98)
qqHH vs TTV (0.89)
qqHH vs ggHH (0.84)
qqHH vs ggH (0.97)
eff*purity
Signal eff
DNN(train): 1.334
DNN(test): 1.192
sig-train
bkg-train
sig-test
bkg-test
(a. u.)
NN
ation
ment, the non-resonant
oduction via vector-boson
nts the unique means to
W±) Higgs coupling (C2V).
arated efficiently from huge
few-observables cut-based
arning algorithm is used to
e signal- or background-like.
Signal to background discrimination for the production of double Higgs boson events via vector boson fusion mechanism in the decay channel with four charged leptons and two b-jets in the final state at the LHC experiment
Gather
1 Signal topology
2 Analysis strategy
5 Conclusions
References
20th International Workshop on Advanced Computing and Analysis Techniques in Physics Research
29 Nov 2021 - 3 Dec 2021, Daejeon (South Korea) + Gather Town (online), Poster 540

5. Conclusions

- Despite the signal rarity, an area under the ROC curve (AUC) ~ 98% with the DNN algorithm has been computed.
- A similar binary classification task can be performed for discriminating the VBF HH production under Effective Field Theory (EFT) models vs SM bkgs. The former have enhanced cross-sections, and therefore are simpler to be selected.
- The use of Deep Learning techniques will be proposed to Ph.D. students in the 2nd edition of the ML_INFN hackathon (13-15 December 2021).

Signal/background discrimination for the VBF Higgs four lepton decay channel with the CMS experiment : introduction to the exercise

4. Results

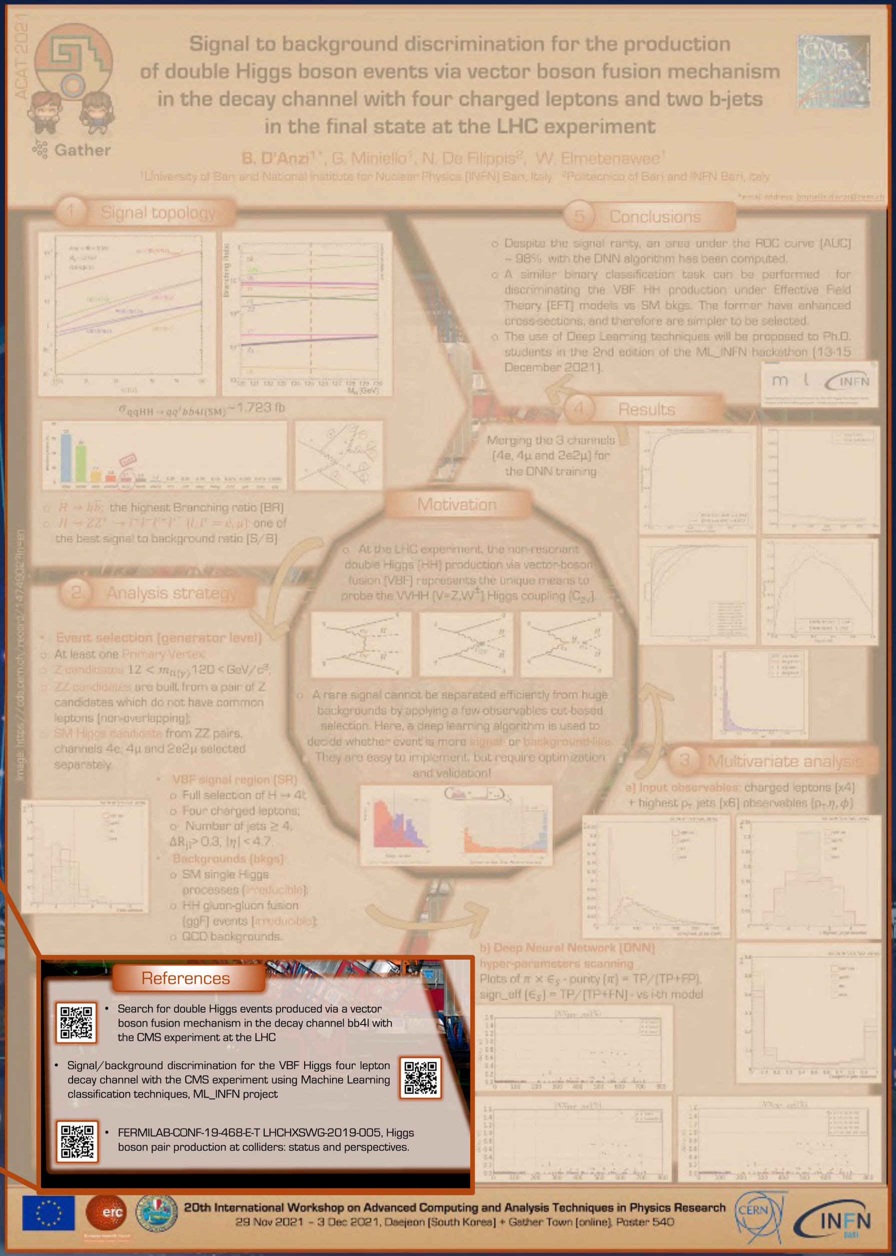


Thank you!

brunella.d’anzi@cern.ch

References

- Search for double Higgs events produced via a vector boson fusion mechanism in the decay channel bb4l with the CMS experiment at the LHC
- Signal/background discrimination for the VBF Higgs four lepton decay channel with the CMS experiment using Machine Learning classification techniques, ML_INFN project
- FERMILAB-CONF-19-468-E-T LHCHXSWG-2019-005, Higgs boson pair production at colliders: status and perspectives.



ACAT 2021

29 NOV - 03 DEC, 2021

AI Decoded:

Towards Sustainable, Diverse, Performant and Effective Scientific Computing

Backup slides

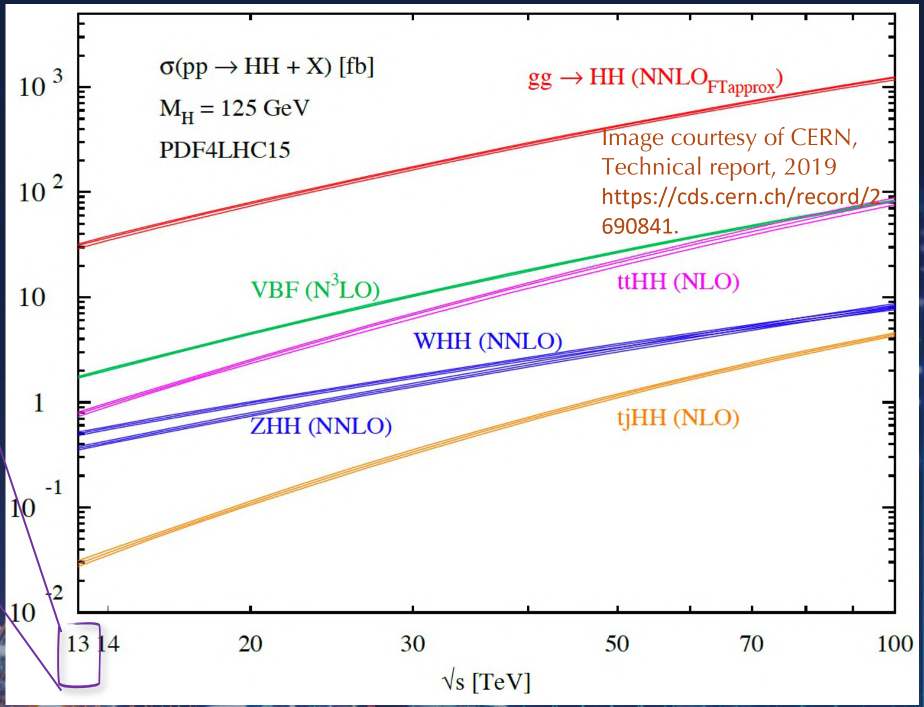
Daejeon, Republic of Korea

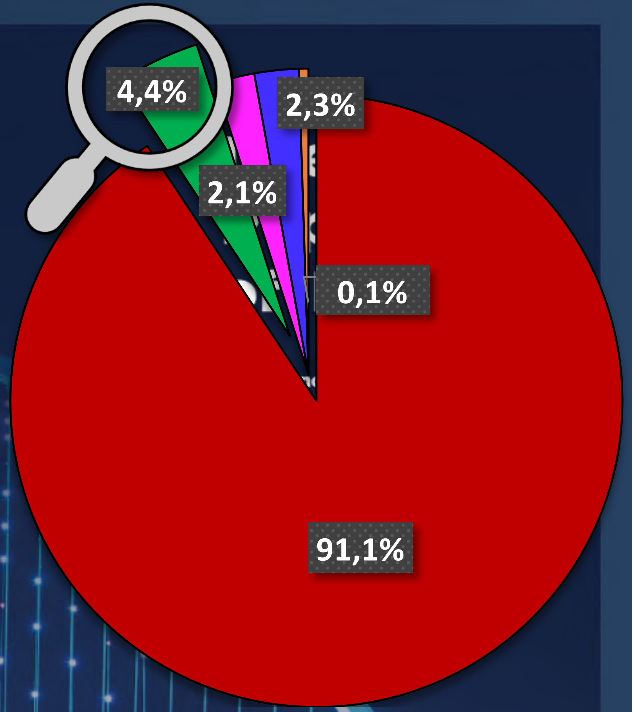
Virtual and
IBS Science Culture Center

Motivation: the HH production mode

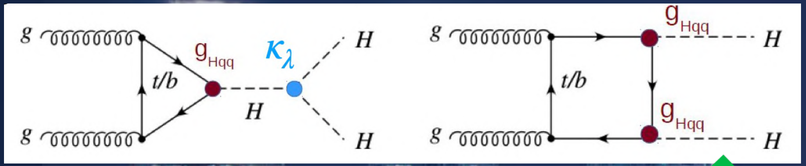
Production mode	σ[fb]
	$\sqrt{s}=13\,\text{TeV}$
Gluon fusion	$33.49^{+4.3\%}_{-6.0\%}$ (scale) $\pm 2.1\%$(PDF) $\pm 2.3\%(\alpha_s)$ $\pm 5.0\%$(top)
VBF	$1.62^{+2.3\%}_{-2.7\%}$ (scale) $\pm 2.3\%$(PDF + α_s)
$t\bar{t}$HH	$0.772^{+1.7\%}_{-4.5\%}$ (scale) $\pm 3.2\%$(PDF + α_s)
W^+HH	$0.329^{+0.32\%}_{-0.41\%}$ (scale) $\pm 2.2\%$(PDF + α_s)
W^-HH	$0.173^{+1.2\%}_{-1.3\%}$ (scale) $\pm 2.8\%$(PDF + α_s)
ZHH	$0.362^{+3.4\%}_{-2.6\%}$ (scale) $\pm 1.9\%$(PDF + α_s)
tjHH	$0.0281^{+5.2\%}_{-3.2\%}$ (scale) $\pm 4.5\%$(PDF + α_s)

LHC center-of-mass energy!

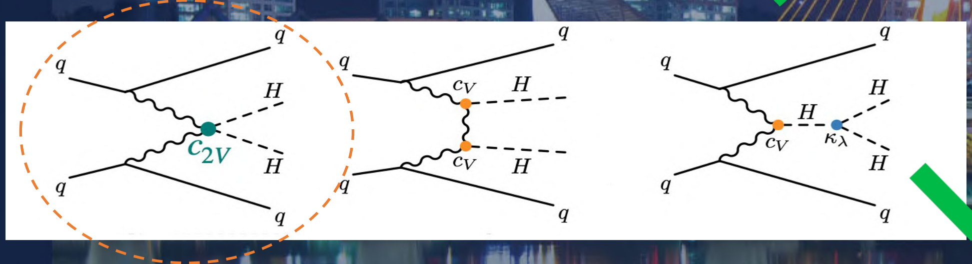


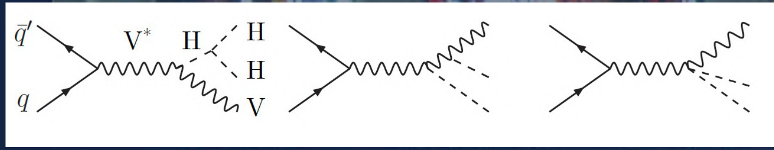


Gluon-gluon fusion (ggF)



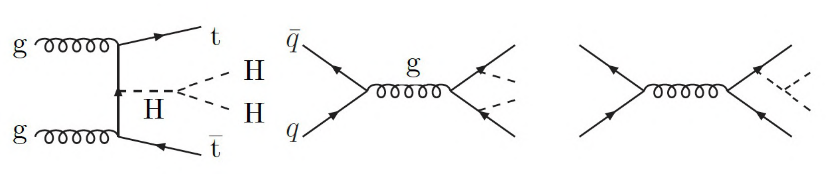
Vector boson fusion (VBF)



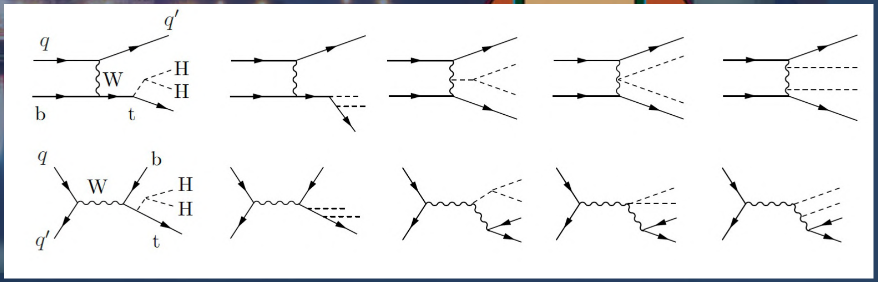


Double Higgs-strahlung/Vector boson associated production (VHH, $V = Z, W^+, W^-$)

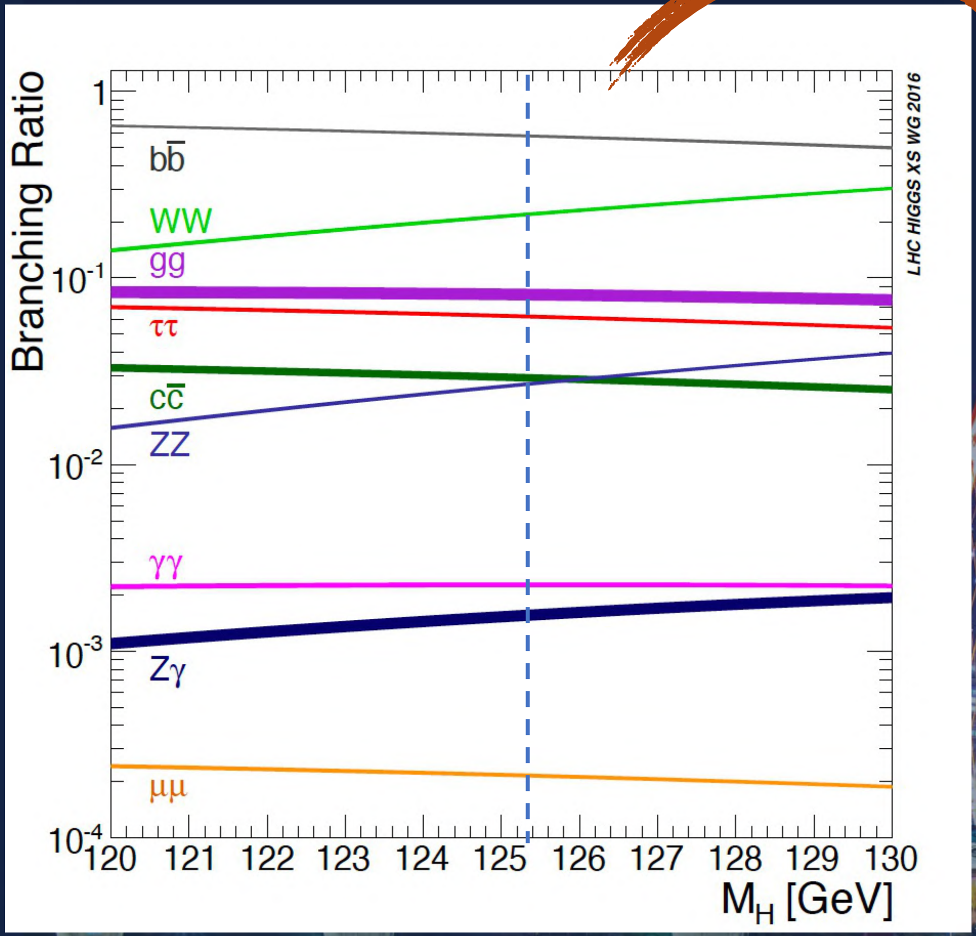
Top quark pair associated production/ Double Higgs bremsstrahlung off top quarks ($t\bar{t}$HH)



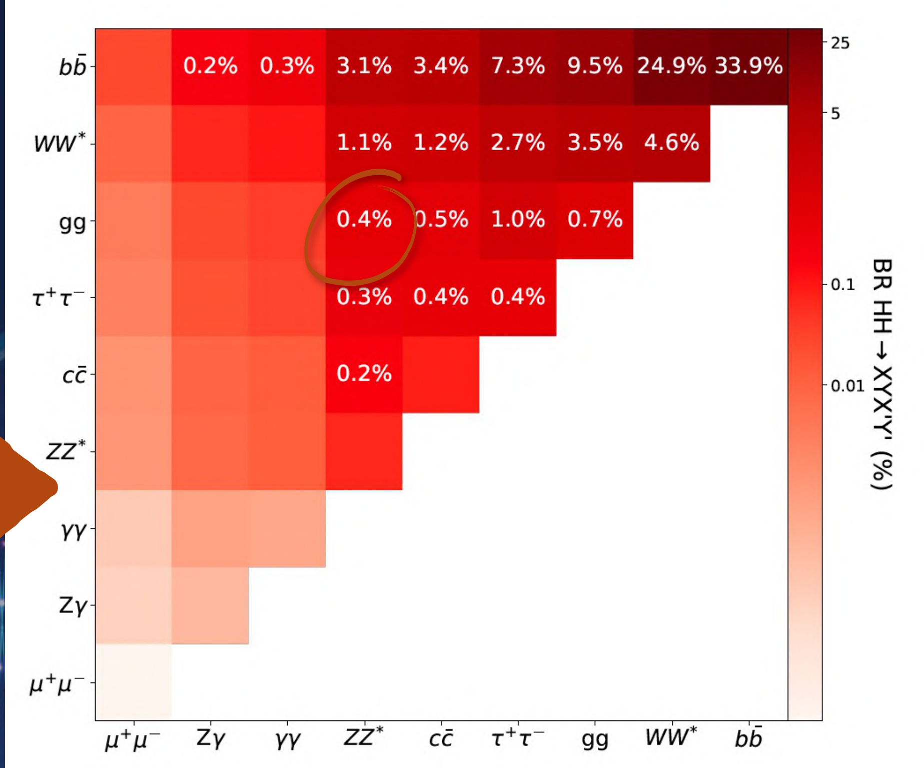
Single top quark associated production (tjHH)

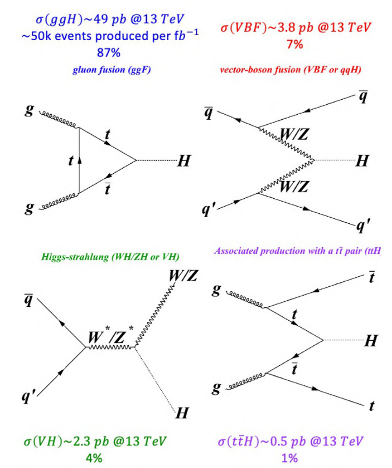


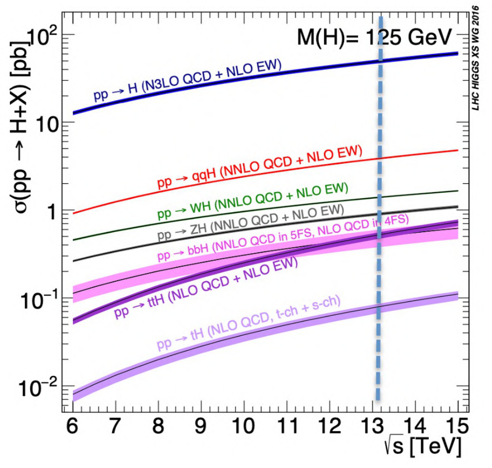
Motivation: the HH decay mode



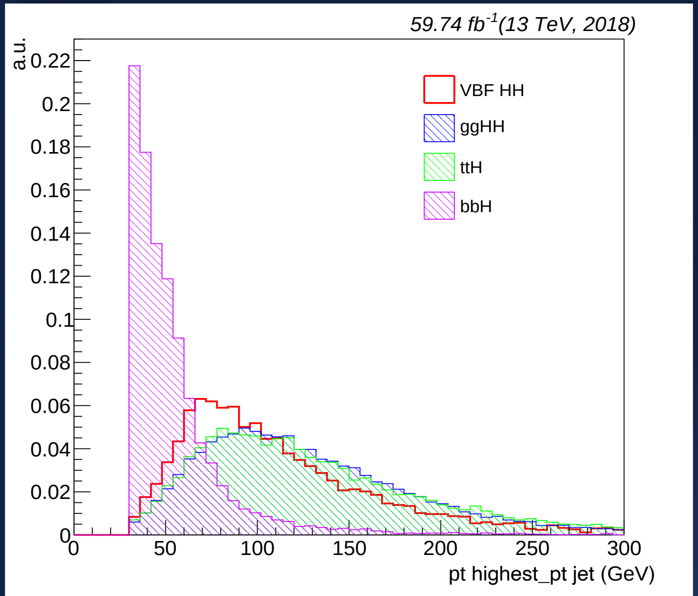
Decay mode	$\mathcal{B}$ [%]
$H \to b\bar{b}$	$58.09^{+0.72}_{-0.73}$
$H \to W^{\pm}W^{\mp *}$	$21.52^{+0.33}_{-0.33}$
$H \to gg$	$8.18^{+0.42}_{-0.42}$
$H \to \tau^+\tau^-$	$6.27^{+0.10}_{-0.10}$
$H \to c\bar{c}$	$2.88^{+0.16}_{-0.06}$
$H \to ZZ^*$	$2.641^{+0.040}_{-0.040}$
$H \to \gamma\gamma$	$0.2270^{+0.0047}_{-0.0047}$
$H \to Z\gamma$	$0.1541^{+0.0090}_{-0.0090}$
$H \to \mu^+\mu^-$	$0.02171^{+0.00036}_{-0.00037}$

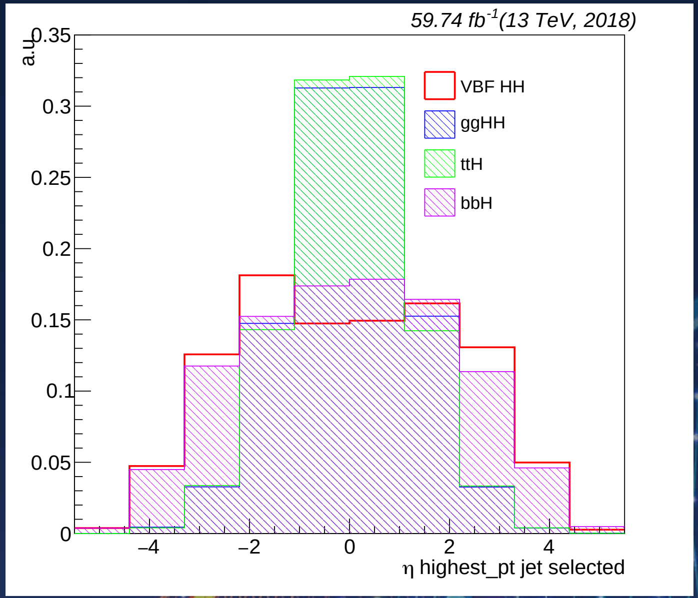


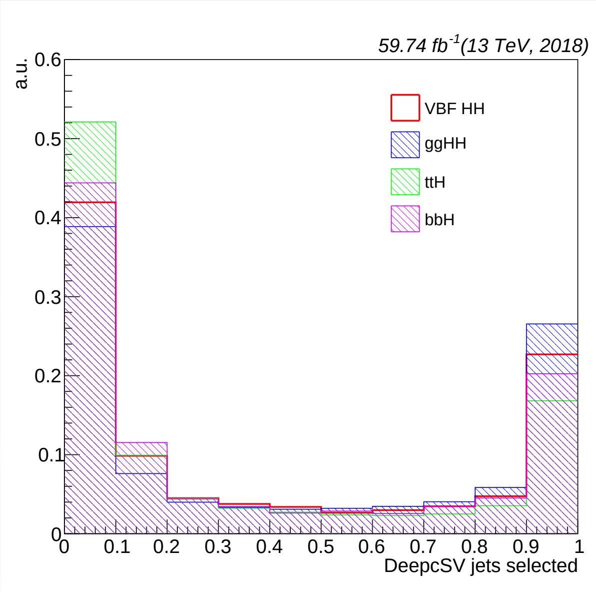




Scanning of the DNN hyper-parameters



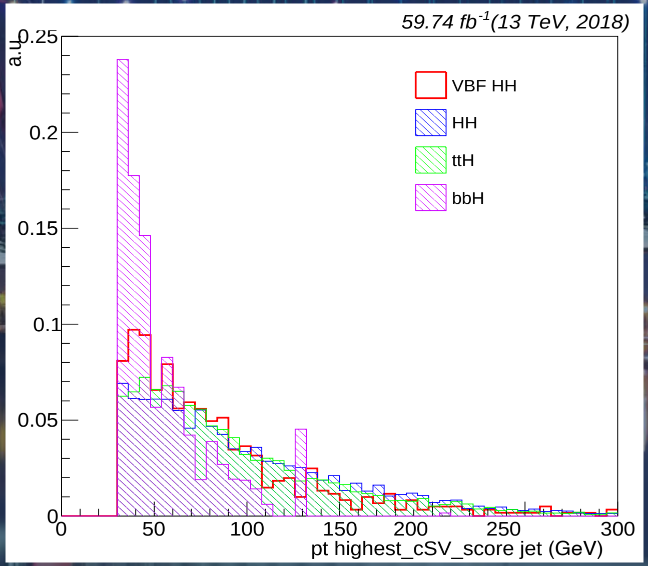




Daejeon, Republic of Korea

Virtual and IBS Science Culture Center

NN hyper-parameters	Tested options
Input variables	leptons/jets (p_T,η,ϕ), jets (Q/G Likelihood, DeepCsV)
Dropout rate	10%, 20%
Topologies	10:10:10:10, 30:30:30:30, 20:20:20:20, 50:50:50:50, ...
Early stop	50, 100, 600, 3000
Minimizer SGD,	SGD, Adadelta
Batch size	5, 32, 64, 128, 786
Neuron	ReLU, SeLU, Tanh
Loss Scaling	MC event weight



DNN training strategy

- The channels (4e, 4μ and 2e2μ selected separately) are merged into just a sample for each MC (the channels will be randomized inside the sample);

- Each merged sample are divided into two independent sets:

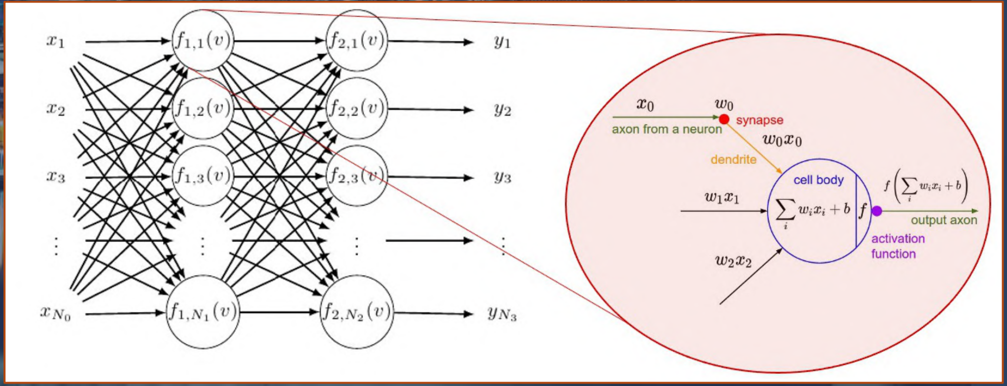
A. **Training set**: contains 80% of all events (from each sample) and is used to train ANNs;
B. **Testing set**: contains remaining 20% of events and is used to test ANN after training;

- ANNs are built and trained via the open-source Keras (standard ML community tool) python package, not linked to TMVA.

Predicted Class

Actual Class		Positive	Negative	
	Positive	True Positive (TP)	False Negative (FN) **Type II Error**	**Sensitivity** $\frac{TP}{(TP+FN)}$
	Negative	False Positive (FP) **Type I Error**	True Negative (TN)	**Specificity** $\frac{TN}{(TN+FP)}$
		Precision $\frac{TP}{(TP+FP)}$	**Negative Predictive Value** $\frac{TN}{(TN+FN)}$	**Accuracy** $\frac{TP+TN}{(TP+TN+FP+FN)}$

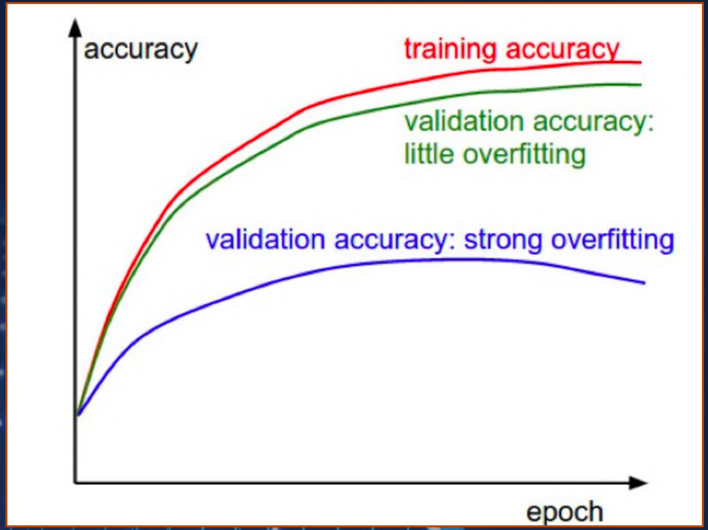
Keras



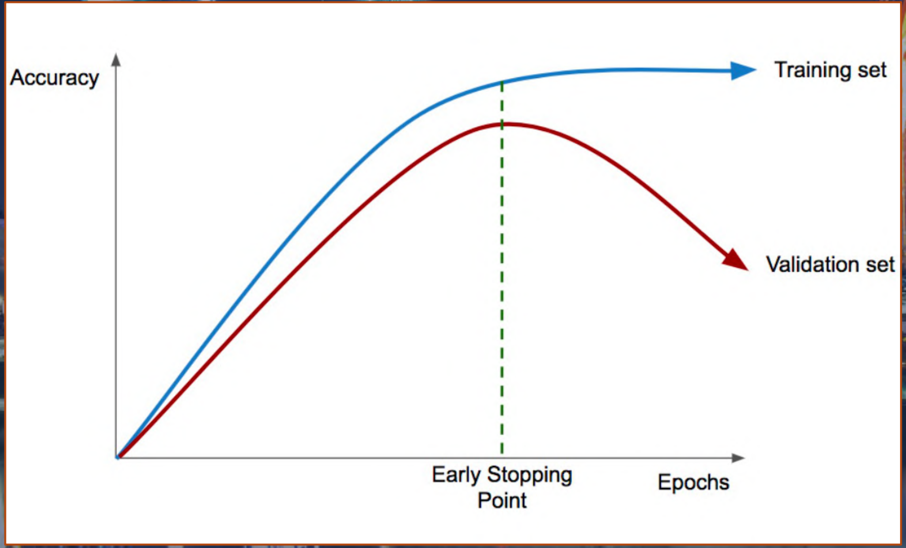
DNN over-training (over-fitting)

Gap between training and validation accuracy indicates the amount of overfitting.

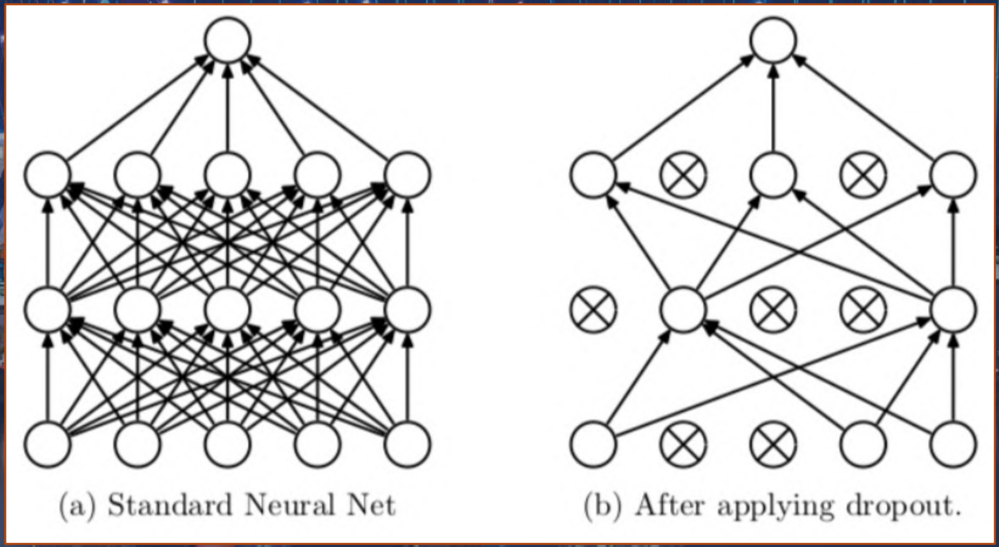
➢ If validation error curve shows small accuracy compared to training it indicates overfitting ⇒ add regularization (**Early Stopping** & **Dropout**) or use more data.



Early Stopping



Dropout



remove nodes randomly during training to avoid co-adaptation on training data