

Chapitre X

Reconstruction de la masse d'une résonance grâce au *Machine Learning*

Sommaire

1	Introduction	1
2	XGBoost	3
2.1	Principe	3
2.2	Entraînement	3
3	DNN	4
3.1	Neurones	4
3.2	Réseaux de neurones	4
3.3	Entraînement	4
4	Événements utilisés	5
4.1	Génération avec FASTSIM	5
4.2	Sélection des événements	6
4.3	Événements obtenus et pondération	7
5	Sélection d'un modèle	7
5.1	Variables d'entrée	7
5.2	Type de modèle	9
5.3	Fonction de coût	9
5.4	Initialisation des poids et optimisation	10
5.5	Structure	11
5.6	Fonction d'activation	13
6	Discussions	13
6.1	Effets de l'intervalle de masse	13
6.2	Effets de l'empilement	13
6.3	Effets de la reconstruction	13
6.4	Effets des faux taus hadroniques	13
6.5	Effets de la séparation des canaux	13
6.6	Effets de bord	13
6.7	Modèle final	13
7	Utilisation du modèle dans les analyses CMS	14
8	Conclusion	14

1 Introduction

AI = exploding for last years

Siri, autonomus cars, AlphaGO, ...

AI > Machine Learning > Deep Learning

principle of ML : get a program mapping inputs to outputs

classification vs regression tasks

in HEP : [1] and the one used for the SM categories in the analysis!

Also cite Gaël thesis :

G. TOUQUET. « Search for an additional neutral MSSM Higgs boson decaying to tau leptons with the CMS experiment ». Thèse de doct. Université Claude Bernard Lyon 1, oct. 2019. URL : <https://hal.archives-ouvertes.fr/tel-02526393>

and Mortiz's :

M. SCHAM. « Standard Model $H \rightarrow \tau\tau$ Analysis with a Neural Network Trained on a Mix of Simulation and Data Samples ». Mém. de mast. Fakultät für Physik des Karlsruher Instituts für Technologie (KIT), juin 2020. URL : <https://publish.etp.kit.edu/record/21993>

and Tanja's :

T. KOPF. « Recoil Calibration as a Neural Network Task ». Mém. de mast. Fakultät für Physik des Karlsruher Instituts für Technologie (KIT), fév. 2019. URL : <https://publish.etp.kit.edu/record/21500>

The Elements of statistical learning : Trees have one aspect that prevents them from being the ideal tool for predictive learning, namely inaccuracy. \rightarrow they work great with the data used to create them, but they are not flexible when it comes to classifying new samples...

► **Aim** : reconstruction of di- τ mass.

▷ Already done in :

P. BÄRTSCHI & coll. « Reconstruction of τ lepton pair invariant mass using an artificial neural network ». *Nuclear Instruments and Methods in Physics Research A* **929** (2019), p. 29-33. DOI : [10.1016/j.nima.2019.03.029](https://doi.org/10.1016/j.nima.2019.03.029). URL : <http://www.sciencedirect.com/science/article/pii/S0168900219303377>

P. BALDI, P. SADOWSKI & D. WHITESON. « Enhanced Higgs Boson to $\tau^+\tau^-$ Search with Deep Learning ». *Physical Review Letters* **114**.11 (mar. 2015). DOI : [10.1103/physrevlett.114.111801](https://doi.org/10.1103/physrevlett.114.111801)

► **BUT**

▷ CMS simulated at phase-0 with Delphes \rightarrow we used CMS Fast Simulation (FASTSIM).

▷ No Pile-Up \rightarrow we added it using the 2017 PU profile.

▷ $m_h \in [80; 300]$ GeV per steps of 5 GeV \rightarrow we do from 50 to 800 per steps of 1.

▷ 270 000 training events \rightarrow we have 2 180 992 ($\times 8$).

▷ 100 000 testing events \rightarrow we have 311 504 ($\times 3$).

Try both XGBOOST and DNN and compare, then apply to MSSM HTT analysis

📎 To cite :

- DELPHES 3.4.2 [7, 8]?
- CMS Fast Simulation (FASTSIM) [9-12]
- PYTHIA 8.235 [13]
- FASTJET [14, 15]
- KERAS [16]
- TENSORFLOW [17]
- XGBOOST [18]
- for an example of nn use in HEP
- D. GUEST & coll. « Jet flavor classification in high-energy physics with deep neural networks ». *Physical Review D* **94**.11 (déc. 2016). DOI : [10.1103/physrevd.94.112002](https://doi.org/10.1103/physrevd.94.112002)
- W. SARLE. « Neural Networks and Statistical Models ». 1994. URL : https://people.orie.cornell.edu/davidr/or474/nn_sas.pdf
- P. BÄRTSCHI & coll. « Reconstruction of τ lepton pair invariant mass using an artificial neural network ». *Nuclear Instruments and Methods in Physics Research A* **929** (2019), p. 29-33. DOI : [10.1016/j.nima.2019.03.029](https://doi.org/10.1016/j.nima.2019.03.029). URL : <http://www.sciencedirect.com/science/article/pii/S0168900219303377>
- SVFIT [20]

L. TORTEROTOT, E. AŞILAR & C. BERNET. *Reconstruction of di-tau mass using Machine Learning*. URL : https://github.com/lucastorterotot/DL_for_HTT_mass

2 XGBoost

why? faster a good at challenges (see Colin's slides from *somewhere in the past*)

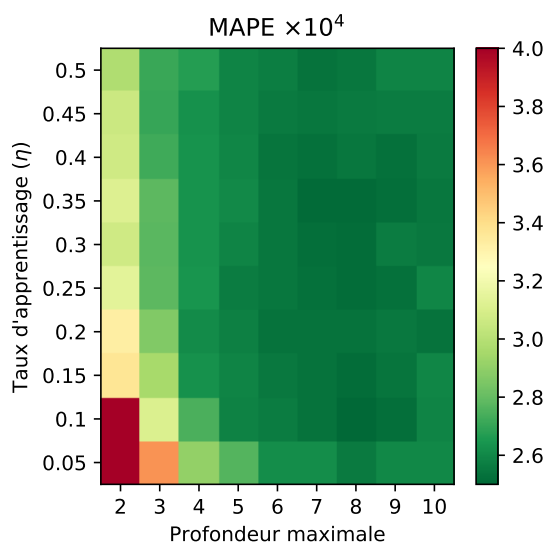
2.1 Principe

tree
max depth, n estimators

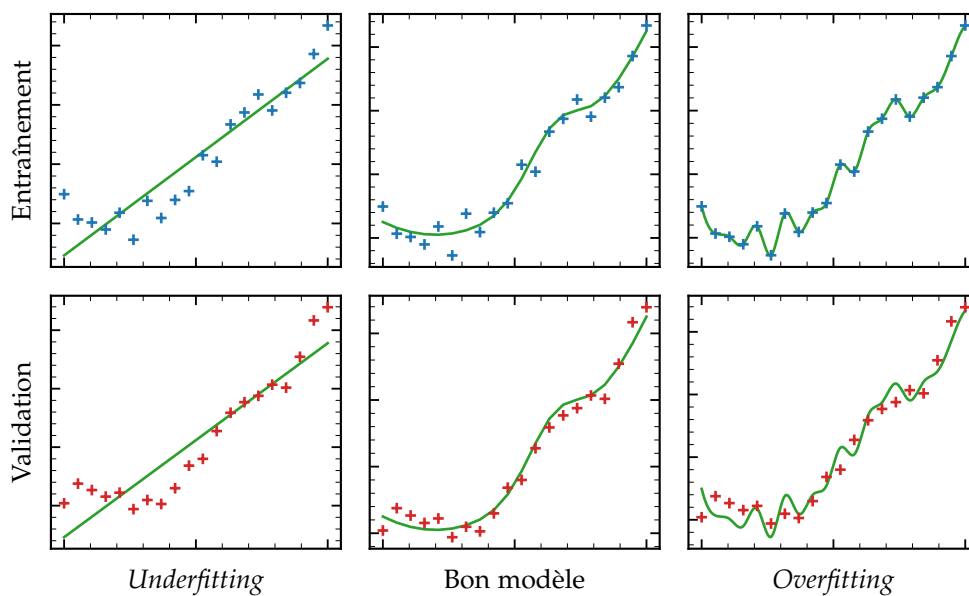
2.2 Entraînement

2.2.1 Amélioration des prédictions

objective
learning rate
early stopping



2.2.2 Surentraînement



3 DNN

I. GOODFELLOW, Y. BENGIO & A. COURVILLE. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016

3.1 Neurones

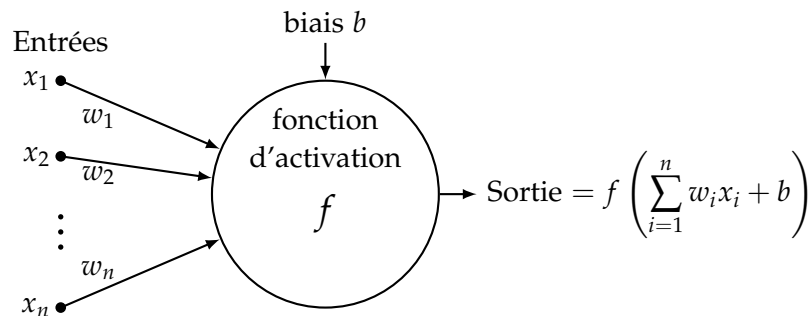


Figure X.1 – Structure d'un neurone. Une fonction f dite d'« activation » est appliquée à la somme des entrées x_i pondérées par les poids w_i et du biais b afin d'obtenir la valeur de sortie.

Activation functions : tanh, sigmoid mostly for classification, linear, relu, elu, selu, softmax, soft-plus ...

3.2 Réseaux de neurones

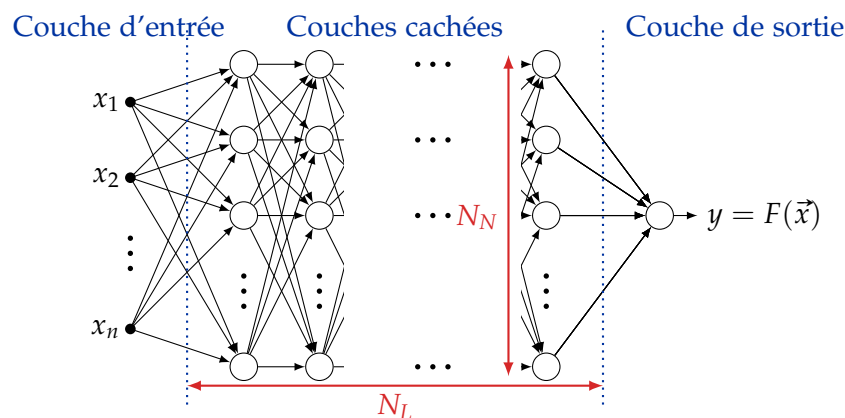


Figure X.2 – Structure d'un réseau de neurones. Une couche d'entrée comporte autant de neurones que de variables x_i . La couche de sortie en comporte autant que de valeurs à donner, c'est-à-dire une. Les fonctions d'activation de ces deux couches sont linéaires. Entre elles se trouvent N_L couches cachées, chacune contenant N_N neurones. Diverses fonctions d'activation peuvent être utilisées dans les couches cachées.

3.3 Entraînement

3.3.1 Fonction de coût

loss == objective
= 0 when prediction == truth
minimize it!

3.3.2 Initialisation des poids et optimisation

Adam, Adadelata, SGD

parameters to optimize = weights and biais

need to init : (Glorot) uniform/normal

Glorot : X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks”, in Proceedings of the thirteenth international conference on artificial intelligence and statistics, p. 249. 2010.

4 Événements utilisés

L’objectif des modèles à entraîner est de reconstruire la masse des particules se désintégrant en paire de leptons tau. Il s’agit d’une tâche de régression, il faut donc entraîner les modèles sur le plus de valeurs différentes possible. Dans l’optique d’une utilisation dans les analyses telles que celle présentée dans le chapitre 5, il a été choisi d’utiliser des événements $\mathcal{H} \rightarrow \tau\tau$ où \mathcal{H} est le boson de Higgs du modèle standard h dont la masse est modifiée. La cible du modèle est donc la masse $m_{\mathcal{H}}$.

4.1 Génération avec FASTSIM

Nous avons généré nos propres données simulées [23] afin d’obtenir des événements indépendants de ceux utilisés dans les analyses. Dans le contexte de la collaboration CMS, nous avons utilisé FASTSIM [9-12]. Cet outil permet de procéder à l’ensemble de la simulation des événements introduite chapitre 3, de la génération du processus physique à la reconstruction des objets physiques par le détecteur.

Les processus physiques sont générés par PYTHIA 8 [13] avec les réglages CUEP8M1 [24, 25]. L’énergie dans le centre de masse est de 13 TeV. Pour ne pas générer d’événements indésirables, seules les collisions créant un boson de Higgs par fusion de gluons, mode dominant pour le modèle standard, sont autorisées. De plus, le rapport de branchement $\mathcal{BR}(\mathcal{H} \rightarrow \tau\tau)$ est fixé à 1, c’est-à-dire que \mathcal{H} se désintègre forcément en paires de leptons taus.

La masse de \mathcal{H} varie de 50 à 800 GeV par pas de 1 GeV. Il est important d’utiliser l’intervalle le plus étendu possible, il correspond à la gamme utile des modèles obtenus. L’effet de l’étendue de cet intervalle est discuté dans la section 6. Lorsque $m_{\mathcal{H}}$ est supérieure à 800 GeV, les propriétés de \mathcal{H} , basées sur celles de h , ne permettent pas d’obtenir des valeurs de $m_{\mathcal{H}}$ cohérentes avec la méthode de génération utilisée. Nous ne considérerons pas de masse plus haute. Bien qu’il soit possible pour une particule de se désintégrer en deux taus dès que sa masse est plus élevée que $2m_{\tau} = 3,5$ GeV, la sélection des événements présentée dans la section 4.2 rejette plus de 99 % des événements lorsque $m_{\mathcal{H}} < 50$ GeV. Nous ne considérerons pas de masse plus basse. L’efficacité des sélections appliquées est représentée sur la figure X.3. S’il est possible d’appliquer des poids aux événements afin d’équilibrer l’entraînement sur l’ensemble des valeurs de la cible, plus d’événements sont générés à basse masse afin d’obtenir des topologies d’événements variées malgré la faible efficacité de sélection. Ainsi, la quantité d’événements générés pour chaque valeur de $m_{\mathcal{H}}$ est de :

- 60 000 pour $50 \geq m_{\mathcal{H}} < 300$;
- 20 000 pour $300 \geq m_{\mathcal{H}} < 500$;
- 10 000 pour $500 \geq m_{\mathcal{H}} \geq 800$.

L’empilement est modélisé par superposition du signal $\mathcal{H} \rightarrow \tau\tau$ à des événements dits de « biais minimum » [13]. Il s’agit d’événements pouvant contenir des interactions dures, mais n’activant pas

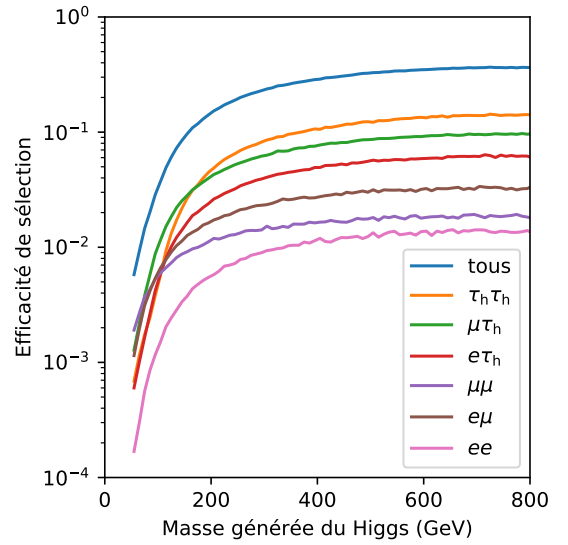


Figure X.3 – Efficacité de sélection des événements pour $m_{\mathcal{H}} \in [50, 800]$ GeV dans les différents canaux et pour tous les canaux.

de chemin de déclenchement. La quantité d'empilement ajoutée à l'événement $\mathcal{H} \rightarrow \tau\tau$ suit le profil de l'année 2017.

4.2 Sélection des événements

4.2.1 Canaux $\tau_h\tau_h$, $\mu\tau_h$, $e\tau_h$ et $e\mu$

La sélection des événements se fait comme exposé dans le chapitre 5 pour l'année 2017 et les canaux $\tau_h\tau_h$, $\mu\tau_h$, $e\tau_h$ et $e\mu$ y étant exploités, à l'exception des coupures servant à séparer la région de signal des régions de contrôle et de détermination, sur $m_T^{(\mu)}$ dans le canal $\mu\tau_h$, $m_T^{(e)}$ dans le canal $e\tau_h$, D_ζ dans le canal $e\mu$. La construction du *dilepton* est inchangée. La correspondance des objets du *dilepton* avec ceux ayant activé le chemin de déclenchement n'est pas vérifiée. Ce choix permet d'obtenir un modèle dont les prédictions auront non seulement un sens dans les régions de contrôle et de détermination, mais aussi plus facilement dans le contexte d'autres analyses dans lesquelles les sélections peuvent différer.

En plus des canaux listés ci-dessus, nous avons également sélectionné des événements des canaux $\mu\mu$ et ee , selon les procédures présentées ci-après.

4.2.2 Canal $\mu\mu$

Sélection des muons Tout muon respectant les critères listés ci-après est retenu pour jouer le rôle de L_1 ou L_2 dans le *dilepton* :

- $p_T^\mu > 10 \text{ GeV}$;
- $|\eta^\mu| < 2.4$;
- paramètres d'impact $d_z < 0,2 \text{ cm}$ et $d_{xy} < 0,045 \text{ cm}$;
- $I^\mu < 0,15 p_T^\mu$;
- passer le point de fonctionnement *medium* du *muonID*.

Sélection du *dilepton* L'événement est retenu à condition qu'au moins une paire $L_1L_2 = \mu\mu$ puisse être construite avec L_1 et L_2 de charges électriques opposées. Il est de plus requis que L_1 et L_2 soient séparés dans le plan (η, ϕ) tel que $\Delta R > 0,3$. Si plus d'une paire possible existe dans l'événement, une seule est retenue selon la logique exposée dans le chapitre 5.

Vétos de leptons supplémentaires Les vetos de leptons supplémentaires doivent être respectés, c'est-à-dire que l'événement ne contient pas :

- de second muon tel que $p_T^\mu > 10 \text{ GeV}$, $|\eta^\mu| < 2,4$, passant le point de fonctionnement *medium* du *muonID* et d'isolation $I^\mu < 0,3 p_T^\mu$;
- de second électron tel que $p_T^e > 10 \text{ GeV}$, $|\eta^e| < 2,5$, passant le point de fonctionnement à 90 % d'efficacité de l'*electron ID MVA* et d'isolation $I^e < 0,3 p_T^e$, l'électron devant passer le veto d'électron de conversion et présenter moins de deux points de passage manquants dans le trajectographe.

4.2.3 Canal ee

Sélection des électrons Tout électron respectant les critères listés ci-après est retenu pour jouer le rôle de L_1 ou L_2 dans le *dilepton* :

- $p_T^e > 20 \text{ GeV}$;
- $|\eta^e| < 2.4$;
- paramètres d'impact $d_z < 0,2 \text{ cm}$ et $d_{xy} < 0,045 \text{ cm}$;
- $I^e < 0,1 p_T^e$;
- passer le point de fonctionnement à 90 % d'efficacité de l'*electron ID MVA*.

Sélection du dilepton L'événement est retenu à condition qu'au moins une paire $L_1 L_2 = ee$ puisse être construite avec L_1 et L_2 de charges électriques opposées. Il est de plus requis que L_1 et L_2 soient séparés dans le plan (η, ϕ) tel que $\Delta R > 0,5$. Si plus d'une paire possible existe dans l'événement, une seule est retenue selon la logique exposée dans le chapitre 5.

Vétos de leptons supplémentaires Les vetos de leptons supplémentaires doivent être respectés, c'est-à-dire que l'événement ne contient pas:

- de second muon tel que $p_T^\mu > 10 \text{ GeV}$, $|\eta^\mu| < 2,4$, passant le point de fonctionnement *medium* du *muonID* et d'isolation $I^\mu < 0,3 p_T^\mu$;
- de second électron tel que $p_T^e > 10 \text{ GeV}$, $|\eta^e| < 2,5$, passant le point de fonctionnement à 90 % d'efficacité de l'*electron ID MVA* et d'isolation $I^e < 0,3 p_T^e$, l'électron devant passer le veto d'électron de conversion et présenter moins de deux points de passage manquants dans le trajectographe.

4.3 Événements obtenus et pondération

Plus de 22 millions d'événements ont été générés. Environ 3 millions sont sélectionnés selon les critères présentés précédemment. La distribution de $m_{\mathcal{H}}$ dans ces événements sélectionnés est représentée sur la figure X.4a. Quelques événements présentent des valeurs de $m_{\mathcal{H}}$ au-delà de 800 GeV, cet effet est dû à la largeur de cette particule. À haute masse, la durée de vie est réduite. Alors, l'incertitude sur la masse augmente, comme l'indique le principe d'incertitude de Heisenberg. Cet effet n'est pas présent à basse masse. Les événements retenus dans la suite sont ceux où la masse effective de \mathcal{H} se situe bien entre 50 et 800 GeV.

Ces événements sont de plus séparés en trois groupes selon les proportions suivantes :

- 70 % pour l'entraînement. Ce sont ces événements que les modèles pourront exploiter afin d'apprendre à prédire correctement $m_{\mathcal{H}}$;
- 20 % pour la validation. Ces événements permettent de vérifier qu'il n'y a pas de surentraînement, c'est-à-dire que le modèle ne se spécialise pas vis-à-vis du jeu d'entraînement;
- 10 % pour les tests. Ces événements ne sont pas utilisés lors des entraînements et permettent donc de tester les modèles sur des données inédites. Sauf contre-indication, les figures sont toutes obtenues avec ce groupe d'événements.

Afin de réaliser un entraînement équitable entre les différentes valeurs de $m_{\mathcal{H}}$, un poids est associé à chaque événement de manière à ce que la distribution pondérée de $m_{\mathcal{H}}$ soit plate. La distribution pondérée de $m_{\mathcal{H}}$ sur les événements utilisés pour l'entraînement des modèles est représentée sur la figure X.4b.

5 Sélection d'un modèle

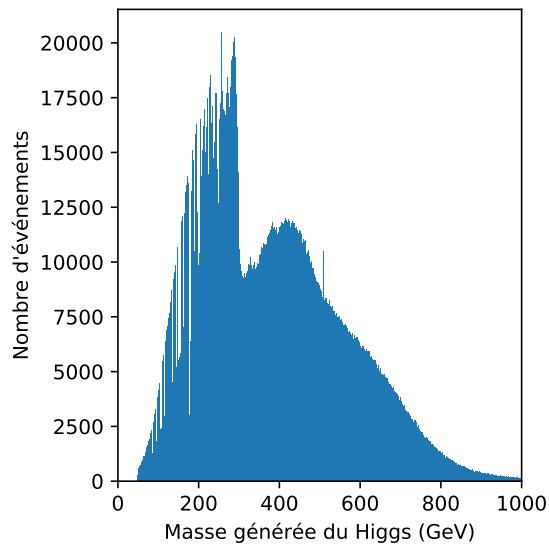
no customized loss yet.

hard to get one single score to determine which model is the best : use mse, mae, mape, median diff, ± 1 or 2σ width ... low, medium, high and full mass regions as well.

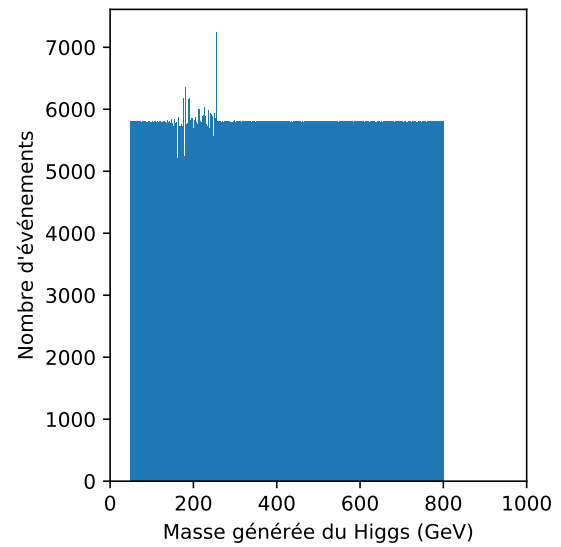
5.1 Variables d'entrée

Model inputs : DNN not that sensible but XGB is better when having all of them, then use all inputs (give list) and not a subset of them.

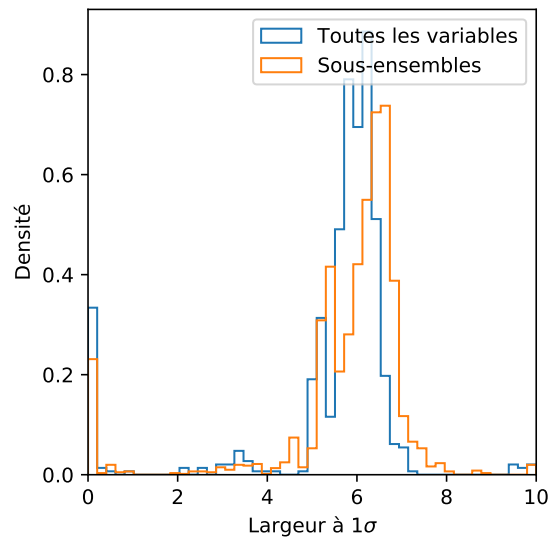
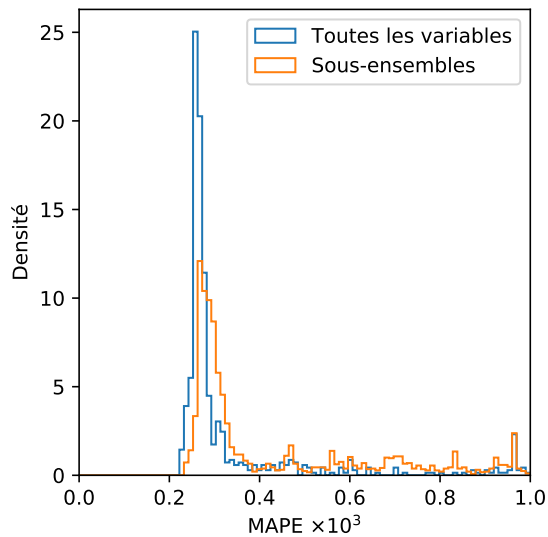
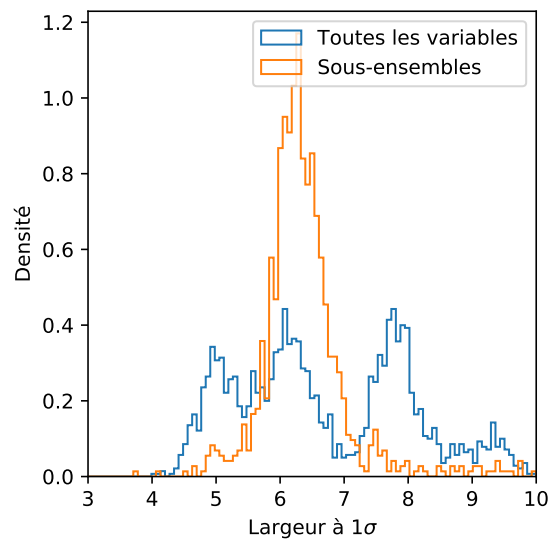
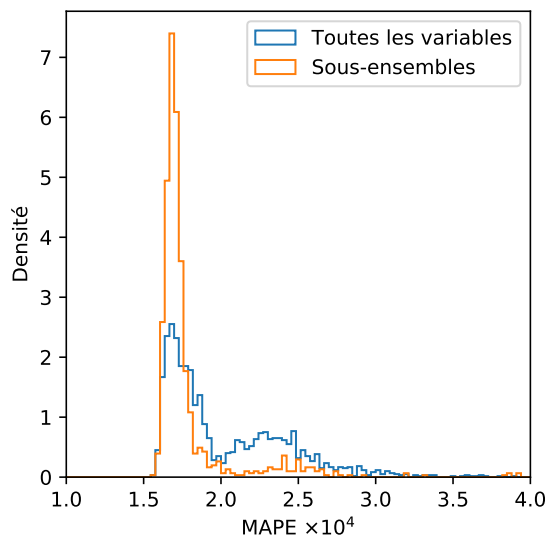
use plots with ref when relevant



(a) Distribution brute sur tous les événements.

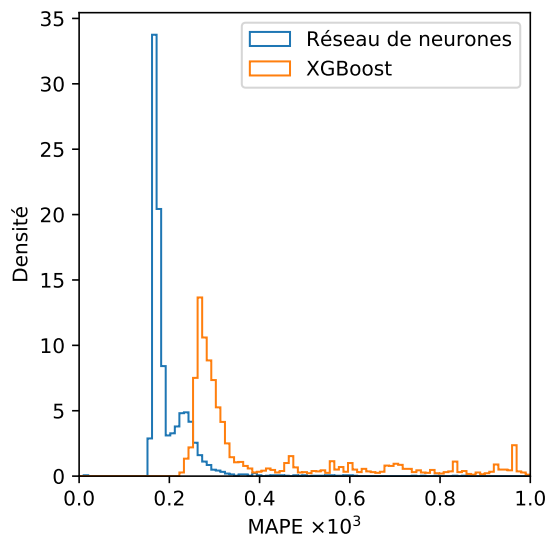
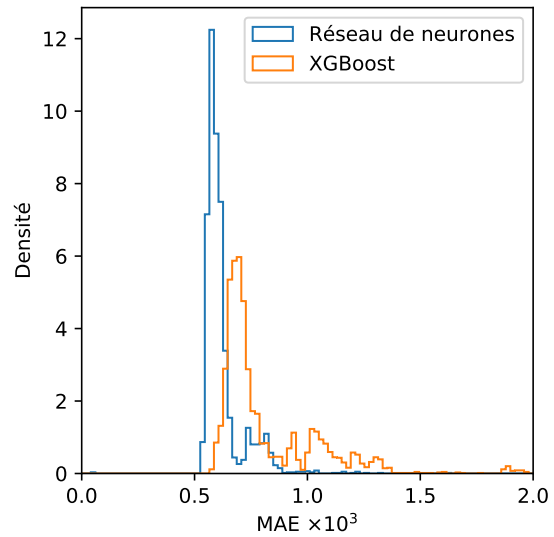
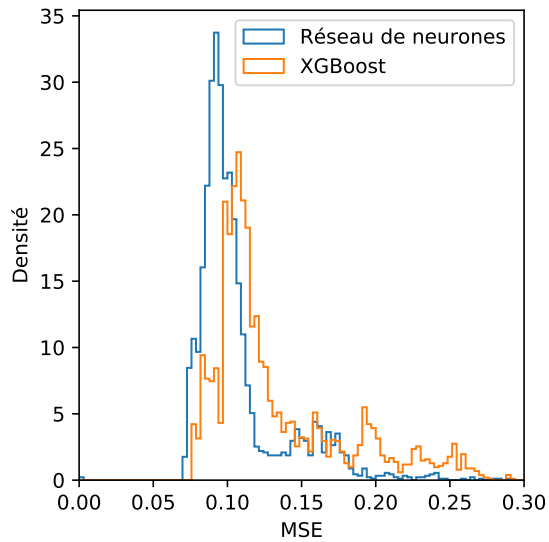


(b) Distribution pondérée pour les événements d'entraînement.

Figure X.4 – Distributions de la masse générée de \mathcal{H} .

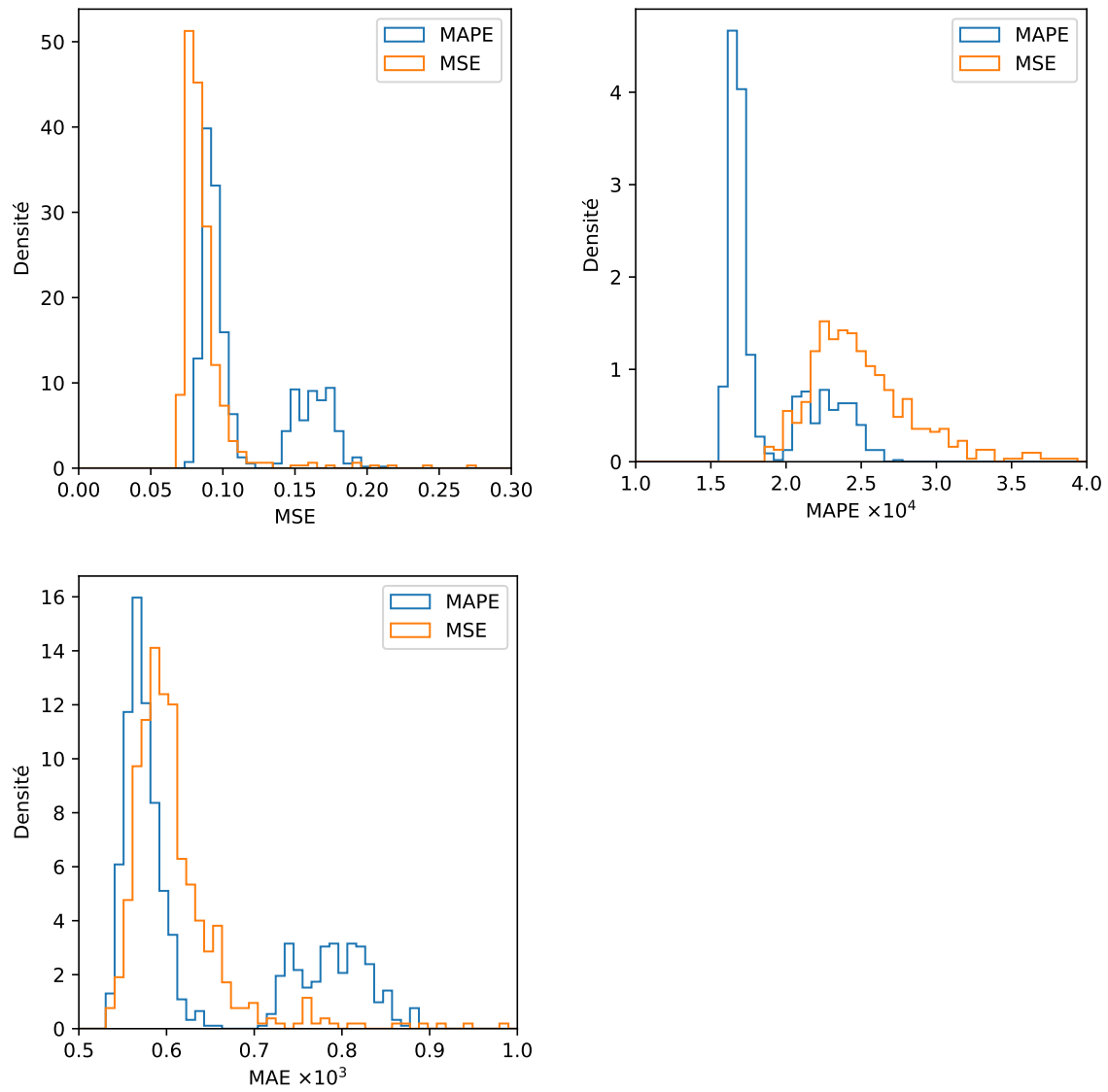
5.2 Type de modèle

DNN vs XGB : use DNN!



5.3 Fonction de coût

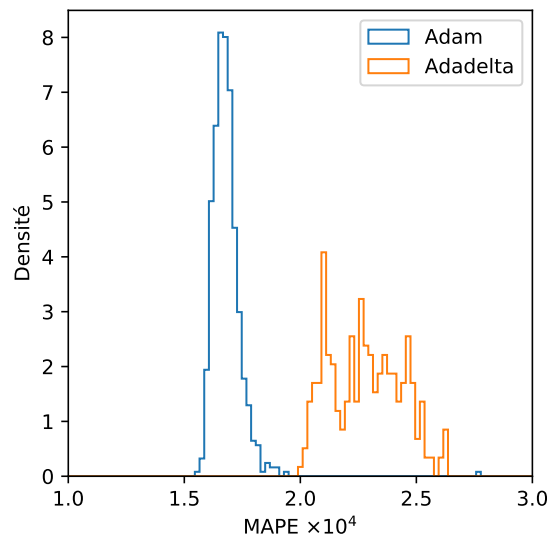
loss : when using a given loss, the corresponding models are of course better when using the loss as score.



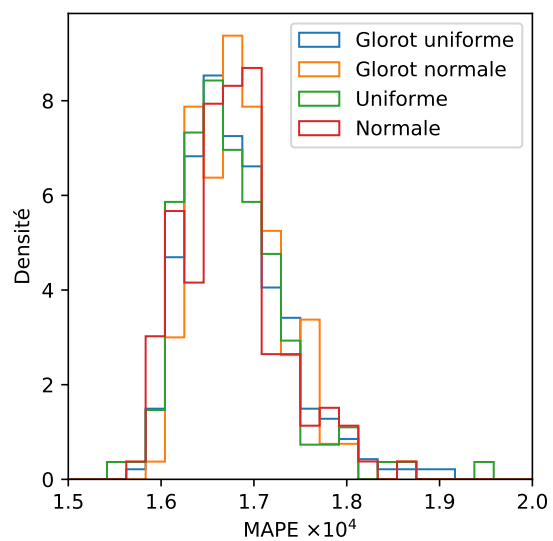
use mape loss, gives the better results

5.4 Initialisation des poids et optimisation

optimizer : Adam

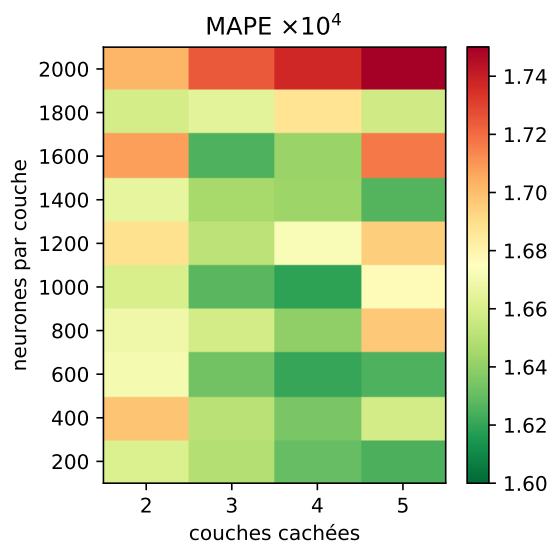


w init mode Glorot Uniform

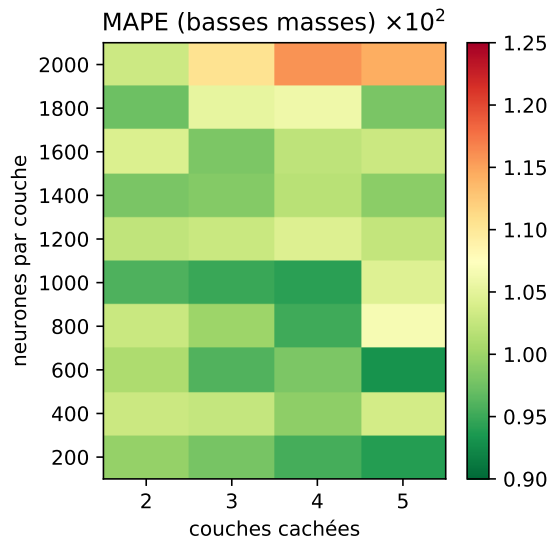


5.5 Structure

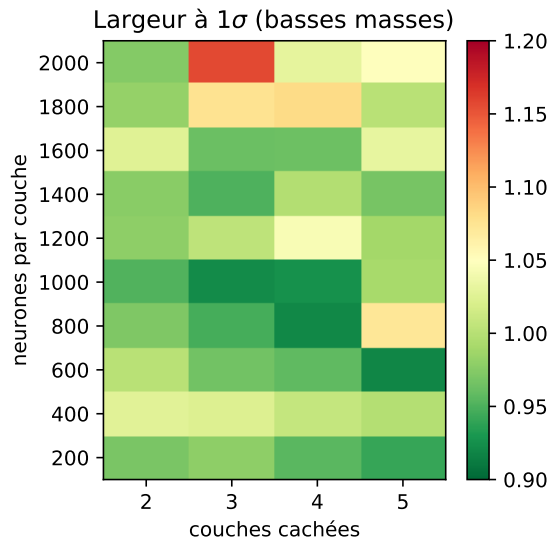
which structure?



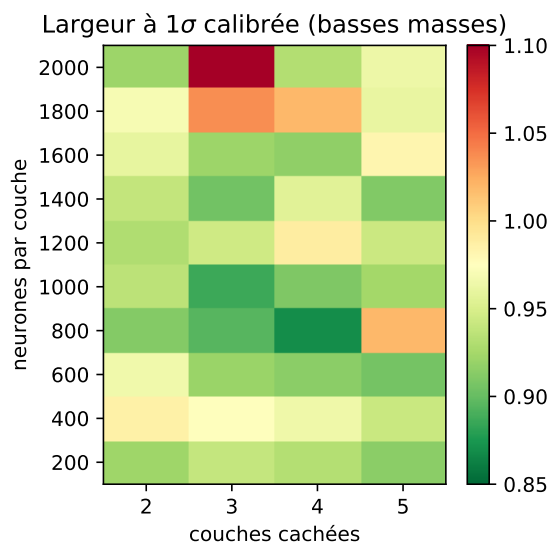
several possibilities, but the loss mass region contains the Z boson and is important



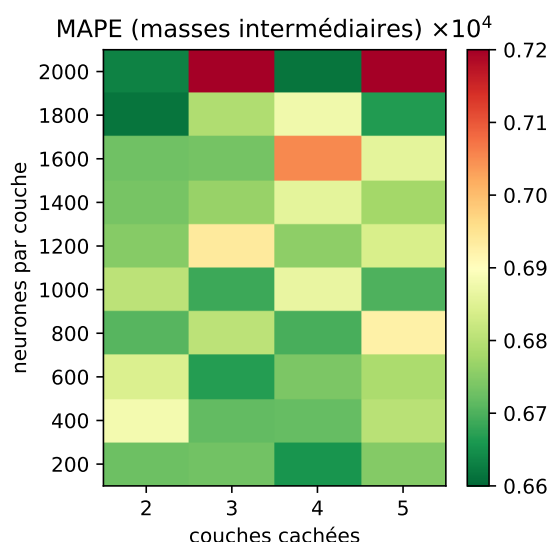
2x900 and 5x600 seem to be the best options, check the low mass resolution



5x600 seems good, check the low mass **calibrated** resolution



and in the medium mass region we have



3x1000 is the best compromise we found

5.6 Fonction d'activation

activation = softplus

6 Discussions

6.1 Effets de l'intervalle de masse

6.2 Effets de l'empilement

also show PU effect (see fig 2 and 3 from report 2020-11-20, update with new models and samples)

6.3 Effets de la reconstruction

show trained/tested on gen tau, gen tau decays, reco tau decays (=real), see fig 3 from report 2021-01-11

the model understand the physics, now it has to deal with the reco resolution and fakes.

6.4 Effets des faux taus hadroniques

6.5 Effets de la séparation des canaux

not relevant (fig3 report 2021-01-21)

6.6 Effets de bord

use the custom loss with boundaries cuts (basically all the report 2021-02-04)

Follow report from 2021-02-04 but for section 3 : We saw that predictions come out too low, which already is a motivation to put larger weights on higher masses, i.e. to weight by truth. Choosing $\sqrt{\text{truth}}$ is of course just a guess then

extend up to 1TeV using the tails

6.7 Modèle final

DEEPTAU

1 TeV

all inputs

activation softplus

```

loss mapesqrt_b
opti Adam
glorot uniform
3 layers of 1000 neurons
show reponses and 2d histo

```

7 Utilisation du modèle dans les analyses CMS

```

show distributions of mTtot and ML predictions
discuss
show limits
discuss

```

8 Conclusion

Références

- [1] D. GUEST & coll. « Jet flavor classification in high-energy physics with deep neural networks ». *Physical Review* **D94.11** (déc. 2016). DOI : [10.1103/physrevd.94.112002](https://doi.org/10.1103/physrevd.94.112002).
- [2] G. TOUQUET. « Search for an additional neutral MSSM Higgs boson decaying to tau leptons with the CMS experiment ». Thèse de doct. Université Claude Bernard Lyon 1, oct. 2019. URL : <https://hal.archives-ouvertes.fr/tel-02526393>.
- [3] M. SCHAM. « Standard Model $H \rightarrow \tau\tau$ Analysis with a Neural Network Trained on a Mix of Simulation and Data Samples ». Mém. de mast. Fakultät für Physik des Karlsruher Instituts für Technologie (KIT), juin 2020. URL : <https://publish.etp.kit.edu/record/21993>.
- [4] T. KOPF. « Recoil Calibration as a Neural Network Task ». Mém. de mast. Fakultät für Physik des Karlsruher Instituts für Technologie (KIT), fév. 2019. URL : <https://publish.etp.kit.edu/record/21500>.
- [5] P. BÄRTSCHI & coll. « Reconstruction of τ lepton pair invariant mass using an artificial neural network ». *Nuclear Instruments and Methods in Physics Research* **A929** (2019), p. 29-33. DOI : [10.1016/j.nima.2019.03.029](https://doi.org/10.1016/j.nima.2019.03.029). URL : <http://www.sciencedirect.com/science/article/pii/S0168900219303377>.
- [6] P. BALDI, P. SADOWSKI & D. WHITESON. « Enhanced Higgs Boson to $\tau^+\tau^-$ Search with Deep Learning ». *Physical Review Letters* **114.11** (mar. 2015). DOI : [10.1103/physrevlett.114.111801](https://doi.org/10.1103/physrevlett.114.111801).
- [7] J. de FAVEREAU & coll. « DELPHES 3 : a modular framework for fast simulation of a generic collider experiment ». *Journal of High Energy Physics* **2** (fév. 2014). DOI : [10.1007/jhep02\(2014\)057](https://doi.org/10.1007/jhep02(2014)057).
- [8] A. MERTENS. « New features in DELPHES 3 ». *Journal of Physics : Conference Series* **608.1** (2015). Sous la dir. de L. FIALA, M. LOKAJICEK & N. TUMOVA. DOI : [10.1088/1742-6596/608/1/012045](https://doi.org/10.1088/1742-6596/608/1/012045).
- [9] S. ABDULLIN & coll. « The Fast Simulation of the CMS Detector at LHC ». *Journal of Physics : Conference Series* **331.3** (déc. 2011). DOI : [10.1088/1742-6596/331/3/032049](https://doi.org/10.1088/1742-6596/331/3/032049).
- [10] A. GIAMMANCO. « The Fast Simulation of the CMS Experiment ». *Journal of Physics : Conference Series* **513.2** (juin 2014). DOI : [10.1088/1742-6596/513/2/022012](https://doi.org/10.1088/1742-6596/513/2/022012).
- [11] M. KOMM. « Fast emulation of track reconstruction in the CMS simulation ». *Journal of Physics : Conference Series* **898** (oct. 2017). DOI : [10.1088/1742-6596/898/4/042034](https://doi.org/10.1088/1742-6596/898/4/042034).
- [12] S. SEKMEN. *Recent Developments in CMS Fast Simulation*. 2017. arXiv : [1701.03850](https://arxiv.org/abs/1701.03850).
- [13] T. SJÖSTRAND & coll. « An Introduction to PYTHIA 8.2 ». *Computer Physics Communications* **191** (2015), p. 159-177. DOI : [10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024). arXiv : [1410.3012](https://arxiv.org/abs/1410.3012) [hep-ph].

- [14] M. CACCIARI, G. P. SALAM & G. SOYEZ. « FASTJET user manual ». *European Physical Journal* **C72** (nov. 2012). DOI : [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2). arXiv : [1111.6097](https://arxiv.org/abs/1111.6097) [hep-ph].
- [15] M. CACCIARI & G. P. SALAM. « Dispelling the N^3 myth for the k_T jet-finder ». *Physics Letters* **B641.1** (sept. 2006), p. 57-61. DOI : [10.1016/j.physletb.2006.08.037](https://doi.org/10.1016/j.physletb.2006.08.037).
- [16] F. CHOLLET & coll. KERAS. <https://keras.io>. 2015.
- [17] M. ABADI & coll. TENSORFLOW : *Large-scale machine learning on heterogeneous distributed systems*. Software available from tensorflow.org. 2015. URL : <https://www.tensorflow.org/>.
- [18] T. CHEN & C. GUESTRIN. « XGBOOST : A Scalable Tree Boosting System ». *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (août 2016). DOI : [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785).
- [19] W. SARLE. « Neural Networks and Statistical Models ». 1994. URL : https://people.orie.cornell.edu/davidr/or474/nn_sas.pdf.
- [20] L. BIANCHINI & coll. « Reconstruction of the Higgs mass in $H \rightarrow \tau\tau$ Events by Dynamical Likelihood techniques ». *Journal of Physics : Conference Series* **513.2** (juin 2014). DOI : [10.1088/1742-6596/513/2/022035](https://doi.org/10.1088/1742-6596/513/2/022035).
- [21] L. TORTEROTOT, E. AŞILAR & C. BERNET. *Reconstruction of di-tau mass using Machine Learning*. URL : https://github.com/lucastorterotot/DL_for_HTT_mass.
- [22] I. GOODFELLOW, Y. BENGIO & A. COURVILLE. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016.
- [23] E. AŞILAR. *How to produce nanoAOD events of $h \rightarrow \tau\tau$ where Higgs has a 130 GeV mass*. URL : https://github.com/easilar/cmssw/blob/from-CMSSW_10_2_22/README.
- [24] The CMS Collaboration. « Event generator tunes obtained from underlying event and multiparton scattering measurements ». *European Physical Journal* **C76.3** (2016). DOI : [10.1140/epjc/s10052-016-3988-x](https://doi.org/10.1140/epjc/s10052-016-3988-x). arXiv : [1512.00815](https://arxiv.org/abs/1512.00815) [hep-ex].
- [25] The CMS Collaboration. « Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements ». *European Physical Journal* **C80** (mar. 2019). DOI : [10.1140/epjc/s10052-019-7499-4](https://doi.org/10.1140/epjc/s10052-019-7499-4). URL : <https://cds.cern.ch/record/2669320>.

Table des matières

X	Reconstruction de la masse d'une résonance grâce au <i>Machine Learning</i>	1
1	Introduction	1
2	XGBOOST	3
	2.1 Principe	3
	2.2 Entraînement	3
3	DNN	4
	3.1 Neurones	4
	3.2 Réseaux de neurones	4
	3.3 Entraînement	4
4	Événements utilisés	5
	4.1 Génération avec FASTSIM	5
	4.2 Sélection des événements	6
	4.3 Événements obtenus et pondération	7
5	Sélection d'un modèle	7
	5.1 Variables d'entrée	7
	5.2 Type de modèle	9
	5.3 Fonction de coût	9
	5.4 Initialisation des poids et optimisation	10
	5.5 Structure	11
	5.6 Fonction d'activation	13
6	Discussions	13
	6.1 Effets de l'intervalle de masse	13
	6.2 Effets de l'empilement	13
	6.3 Effets de la reconstruction	13
	6.4 Effets des faux taus hadroniques	13
	6.5 Effets de la séparation des canaux	13
	6.6 Effets de bord	13
	6.7 Modèle final	13
7	Utilisation du modèle dans les analyses CMS	14
8	Conclusion	14

