

Machine learning dans ATLAS

Yann Coadou

CPPM Marseille

BigData au CPPM
CPPM, 16 October 2017





- Physique au LHC : très souvent un faible signal et un large bruit de fond
- Exemple au CPPM : analyse $t\bar{t}H(\rightarrow b\bar{b})$ dans ATLAS
- Techniques avancées d'analyse multivariée indispensables pour diminuer la quantité de données nécessaire pour observer un phénomène
- De nombreuses publications ATLAS/CMS/LHCb ne seraient pas assez sensibles sans ces techniques
⇒ moins de résultats de physique, moins d'informations fournies aux théoriciens pour étendre le modèle standard



- Courant dans ATLAS (et CMS/LHCb) : arbres de décision boostés (BDT), réseaux de neurones à une seule couche cachée
- Utilisé de deux façons : classification (signal contre bruit de fond) et régression (estimation d'une valeur)
- La plupart du temps à partir des logiciels TMVA et NeuroBayes
- Dans ATLAS : de nombreuses analyses (dont $t\bar{t}H$ au CPPM), b -tagging (CPPM, d'abord ANN puis BDT)
- Parfois poussé à l'extrême. Exemple CMS $H \rightarrow \gamma\gamma$ ▶ CMS-PAS-HIG-13-001 :
 - vertex sélectionné par BDT
 - 2^e vertex BDT pour estimer probabilité d'être à moins d'1 cm du point d'interaction
 - identification photon avec BDT
 - énergie des photons corrigée avec régression BDT
 - incertitude en énergie événement par événement par BDT
 - plusieurs BDT pour extraire signal dans différentes catégories
- Petit nouveau en physique des particules : deep learning (apprentissage profond)



- Remplacer les variables discriminantes inspirées par la physique avec des infos plus brutes, comme les quadrivecteurs [▶ arXiv 1402.4735](#)
[▶ arXiv 1410.3469](#)
- Classification des jets de particules : transformer le calorimètre en image et utiliser des convolutional neural networks pour les analyser
[▶ arXiv 1511.05190](#)
- Etiquetage des jets b : utiliser une structure de réseau qui s'adapte au nombre variable de traces dans chaque jet (recurrent neural network [▶ ATL-PHYS-PUB-2017-003](#)), introduction du deep learning ()
- Pour le moment : pas directement utilisé dans une analyse, mais fait son chemin



- **NOTE** : le but n'est pas la R&D sur le deep learning, mais son utilisation optimale dans les analyses de physique (c'est un outil)
- Stage de M2 en 2016 (Ziyu GUO) pour tester le deep learning sur $t\bar{t}H(\rightarrow b\bar{b})$:
 - utilisait exactement la même infrastructure que le BDT déjà développé
 - résultats encourageants ou décevants : performance équivalente pour reconstruire l'état final (avec 6 jets dont 4 jets b) et pour séparer signal et bruit de fond
- Thèse en cours (Ziyu), en collaboration avec le Laboratoire d'Informatique Fondamentale de Marseille (LIF). Perspectives :
 - utilisation de variables moins sophistiquées
 - adaptation aux inputs de taille variable
 - structures de réseaux ou d'entraînement complexes (multi-instance learning, réseaux récurrents, adversarial networks)
 - mettre potentiellement en défaut la technique du "divide and conquer"
 - remplacer les physiciens ?
- Besoin de "veille technologique" pour suivre un domaine bouillonnant