



# Contributions à l'apprentissage et l'inférence adaptatifs – Applications à l'ajustement d'hyperparamètres et à la physique des astroparticules

Rémi Bardenet <sup>1</sup>

---

*Accessit du prix de thèse Gilles Kahn 2013*

*Rémi Bardenet a soutenu sa thèse en novembre 2012 à l'université Paris-Sud XI, sous la direction de Balázs Kégl. Il effectue actuellement un stage post-doctoral au sein du département de statistiques de l'université d'Oxford (Royaume-Uni).*



Les algorithmes d'inférence ou d'optimisation possèdent généralement des hyperparamètres qu'il est nécessaire d'ajuster. Pour une machine à vecteurs supports (SVM), ce sont les paramètres du noyau et le coefficient de la pénalité dans l'objectif de la SVM. Pour un algorithme de Monte Carlo par chaîne de Markov (MCMC), c'est par exemple la distribution de proposition. Ma thèse est consacrée à l'ajustement automatique de ces hyperparamètres, avec des applications en apprentissage artificiel et en physique des particules.

Dans la première moitié de ma thèse, nous avons exploré l'ajustement des hyperparamètres d'algorithmes d'apprentissage artificiel. De récents travaux suggèrent qu'avec les ressources computationnelles dont nous disposons aujourd'hui, l'allocation du temps de calcul devrait favoriser davantage l'exploration de ces hyperparamètres. Une exploration "exhaustive"

---

1. <http://www.2020science.net/people/remi-bardenet>

est toutefois souvent impossible : les réseaux de neurones profonds, qui sont l'état de l'art en classification, possèdent des dizaines d'hyperparamètres. Nous avons d'abord montré que les méthodes dites d'optimisation bayésienne permettent l'ajustement automatique des hyperparamètres de réseaux de neurones profonds. Ensuite, nous avons proposé un algorithme collaboratif d'ajustement, qui mime la mémoire qu'ont les humains d'expériences passées à ajuster le même algorithme sur d'autres données.

La seconde moitié de ma thèse a porté sur les algorithmes Monte Carlo, motivés par l'analyse des données de l'expérience Pierre Auger. Auger est une expérience de physique des particules dédiée à l'observation des gerbes atmosphériques initiées par les rayons cosmiques. Ces gerbes sont de larges cascades de particules élémentaires qui arrosent la surface de la Terre, et qui résultent du choc avec notre atmosphère de noyaux chargés possédant des énergies extrêmement élevées. Il faut en effet imaginer un proton ou un noyau de fer avec l'énergie d'une balle de tennis servie par Maria Sharapova ! Nous avons proposé un modèle statistique du signal d'Auger, dont il s'agissait ensuite d'inférer les paramètres à partir des données. Les algorithmes MCMC sont des algorithmes d'échantillonnage qui explorent des distributions de probabilité complexes. Ils sont un outil de choix pour l'inférence bayésienne dans les modèles complexes comme celui d'Auger, mais leur application n'a pas été triviale.

Notre modèle est une addition de composantes de même type possédant chacune des paramètres différents, chaque composante correspondant à une particule passant dans un détecteur. Dans ce type de modèle, l'inférence bayésienne est sujette à un problème d'identifiabilité des composantes connu sous le nom de *label switching*. Nous avons proposé AMOR, un algorithme MCMC qui ajuste ses hyperparamètres et résout le *label switching*, le tout automatiquement et optimalement. Nous avons prouvé des résultats théoriques de consistance pour une variante d'AMOR, en mettant en lumière des liens intéressants entre le *label switching* et la quantification vectorielle.

Depuis janvier 2013, j'effectue un séjour post-doctoral au département de statistiques de l'Université d'Oxford au Royaume-Uni. Je m'y consacre à développer des méthodes Monte Carlo pour l'inférence bayésienne sur des données massives, avec des applications en analyse de données biologiques.