

## Comparaison de DroQ et TQC

**Mots clés :** IA, apprentissage par renforcement, compromis biais-variance, pytorch

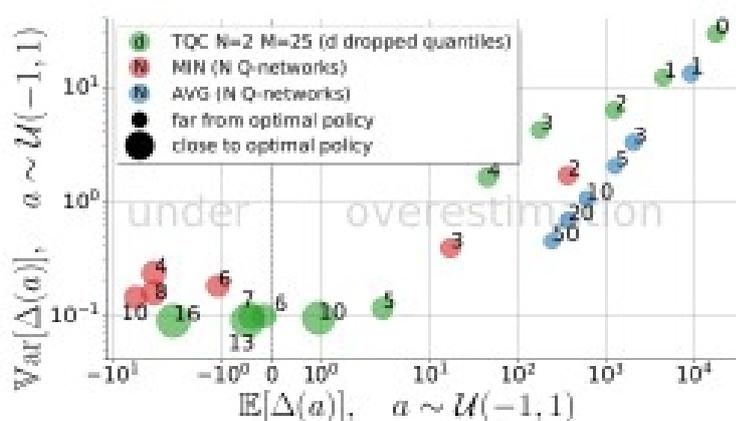
**Encadrant :** Olivier Sigaud, ISIR, Sorbonne Université (Olivier.Sigaud at isir.upmc.fr)

**Nombre d'étudiants :** 3

### Résumé :

L'objectif de ce projet est de comparer deux algorithmes de l'état de l'art en apprentissage par renforcement, DroQ et TQC, en s'intéressant particulièrement à la façon dont ils gèrent le compromis biais-variance.

### Sujet développé :



Les algorithmes DroQ (*Dropout Q-function*) et TQC (*Truncated Quantile Critic*) proposent des avancées intéressantes vis-à-vis de l'état de l'art en apprentissage par renforcement. Partant du même algorithme initial, SAC (*Soft Actor-Critic*), le premier propose d'ajouter des couches de *dropout* et de *layer normalization* dans le critic et de multiplier les critics, tandis que le second propose de multiplier les têtes en sortie de critic et de tronquer les sorties qui semblent les plus sur-estimées.

Ces deux algorithmes n'ont pas été comparés l'un à l'autre. L'objet de ce projet est d'effectuer cette comparaison, en s'intéressant particulièrement à la façon dont ils gèrent le compromis biais-variance.

En pratique, le code de DroQ est fourni par les auteurs, et il existe une version du code de TQC dans la librairie gérée par l'encadrant, BBRL. Après s'être familiarisés avec les algorithmes de l'état de l'art en apprentissage par renforcement en avec BBRL, les étudiants devront développer une version de DroQ dans BBRL, puis réaliser la comparaison attendue en suivant une méthodologie rigoureuse pour laquelle ils seront encadrés.

Pour les étudiants qui seront choisis, les bénéfices de ce projet seront multiples :

- ce sera l'occasion de se familiariser avec les concepts et outils de l'apprentissage par renforcement, qui jouent un rôle significatif dans les progrès actuels de l'Intelligence Artificielle.
- ce sera l'occasion de développer une bonne maîtrise du codage d'algorithmes dans BBRL

- ce sera l'occasion d'apprendre la méthodologie scientifique pour l'évaluation rigoureuse d'algorithmes.

Ce projet faisant appel à de nombreuses notions nouvelles pour les étudiants, une formation initiale à l'apprentissage par renforcement et à des aspects méthodologiques sera assurée durant les premières semaines.

Une forte motivation pour le *machine learning*, des bonnes bases en programmation python et une certaine familiarité avec github sont des prérequis.

## **Références**

L'article « *Controlling Overestimation Bias with Truncated Mixture of Continuous Distributional Quantile Critics* »  
<http://proceedings.mlr.press/v119/kuznetsov20a/kuznetsov20a.pdf>

L'article « *Dropout q-functions for doubly efficient reinforcement learning* »  
<https://arxiv.org/pdf/2110.02034.pdf>

Le dépôt de la librairie BBRL :  
<https://github.com/osigaud/BBRL>