

### Sujet IMPACT Fregier 3

"Réseaux de convolution sur graphes pour la génération de molécules par approche combinée de GANs et apprentissage par renforcement.

La découverte de molécules pour la médecine (médicaments) ou la science des matériaux (stockage d'énergie, cellules photovoltaïques..) est un sujet qui est en train d'être révolutionné par les progrès récents en apprentissage automatique tels que l'apprentissage par renforcement, les réseaux de neurones profonds ou encore les GANs (generative adversarial networks).

En particulier, une approche récente, proposée par des chercheurs d'Harvard, utilise une architecture qui mélange apprentissage par renforcement et GANs. Ce modèle dont l'acronyme est ORGAN se base sur une représentation des molécules sous le format SMILE qui consiste en une chaîne de caractères.

Une autre tendance dans ce domaine est d'essayer de dépasser l'usage du format SMILE et de travailler directement avec une représentation sous forme de graphe. En effet, deux molécules avec des structures spatiales proches ont parfois des représentations SMILE vraiment différentes, donc ce mode de représentation induit des pertes d'informations. Les méthodes qui permettent de travailler directement avec des graphes se basent sur les réseaux de convolutions sur graphes (GCN).

L'objectif de ce projet est de combiner l'approche d'ORGAN (GAN et apprentissage par renforcement) avec les réseaux de convolution sur graphes.

Les compétences nécessaires sont un esprit analytique et un intérêt pour la programmation. Le travail sera fait en python et pytorch et il sera possible de se mettre à niveau au cours du projet.

Le projet sera dirigé par Yaël Frégier, Maître de Conférences à l'Université d'Artois."