

# CHI05 : Chimie et Machine Learning

Septembre 2021 – Avril 2022

---

## COMPOSITION DU GROUPE :

MAALEJ Yassine – GARANCINI Clément - EL-BOUKKOURI Fatima-Zahrae, BARTHES Raphaël

---

## RESUME :

Grâce au Machine Learning et à l'informatique de manière générale, les chimistes ont désormais accès à de nombreux outils pour les accompagner dans leurs recherches. Par ce PSC, notre objectif consiste à perfectionner un outil de Machine Learning en chimie en étudiant les algorithmes existants pour les comprendre et se les approprier afin de les combiner, de les améliorer, et d'implémenter notre propre algorithme de rétrosynthèse.

Nous avons étudié dans un premier temps plusieurs algorithmes open-source de rétrosynthèse pour se faire une idée des méthodes existant déjà. Puis nous nous sommes concentrés sur un algorithme en particulier nommé AiZynthFinder. Nous cherchons ainsi à l'améliorer en nous concentrant seulement sur une partie spécifique de l'algorithme, l'étape de filtration des réactions.

Pour cela, nous avons développé de nombreux modèles de Machine et Deep Learning en reprenant les structures des algorithmes de la littérature. Nous avons entraîné ces modèles sur diverses bases de données, et ce afin d'évaluer leur validité. Nous avons ensuite exploité un logiciel de reconnaissance des groupements chimiques afin de créer nos propres bases de données de réactions spécialisées et d'entraîner par Transfer Learning de nouveaux réseaux dont la mission est de se focaliser sur des groupements ciblés. Les modèles obtenus sont alors légèrement meilleurs que les modèles initiaux. Il s'agit par la suite de tenter d'appliquer cette même méthode à d'autres parties de l'algorithme et de tenter de réduire le temps d'exécution d'un tel algorithme.