



Introduction aux méthodes d'apprentissage en chimie.

François-Xavier COUDERT^a

^aChimie ParisTech, PSL Research University, CNRS,
Institut de Recherche de Chimie Paris, 75005 Paris, France

La chimie est une discipline scientifique qui étudie la composition, la structure et les propriétés des matériaux, ainsi que les réactions chimiques qui ont lieu entre eux. Pour comprendre et prédire ces phénomènes, il est nécessaire de mettre en œuvre des méthodes d'apprentissage automatique, également appelées "*machine learning*" en anglais.

Lors de cette conférence, nous allons présenter les différentes méthodes d'apprentissage automatique qui peuvent être utilisées en chimie, ainsi que leurs avantages et limites. Nous verrons comment ces méthodes peuvent être utilisées pour résoudre des problèmes tels que la classification de molécules, la prédiction de propriétés physiques ou la découverte de nouvelles substances.

Nous aborderons également les différents types de données qui sont couramment utilisées en chimie, ainsi que les méthodes de préparation et de traitement de ces données. Enfin, nous discuterons des enjeux éthiques et sociaux liés à l'utilisation de ces techniques en chimie, et des défis à relever pour leur déploiement à grande échelle.

Références :

- N. Artrith, K. T. Butler, F.-X. Coudert, S. Han, O. Isayev, A. Jain and A. Walsh, "Best practices in machine learning for chemistry", *Nature Chemistry*, 2021, 13 (6), 505–508.
- S. Chibani and F.-X. Coudert, "Machine learning approaches for the prediction of materials properties", *APL Materials*, 2020, 8 (8), 080701.
- Chloé-Agathe Azencott, *Introduction au Machine Learning*, 2^e édition, Dunod, 2022.

Mots Clés : Apprentissage statistique, Bases de données.