

# Présentation de l'ES APPRENTISSAGE ARTIFICIEL (« Machine-Learning »)

Dr Chloé AZENCOTT, Centre de BioInformatique  
& Pr Fabien MOUTARDE, Centre de Robotique  
MINES ParisTech  
PSL Université Paris

Fabien.Moutarde@mines-paristech.fr  
<http://people.mines-paristech.fr/fabien.moutarde>

## Objectifs du cours et intervenants

- **Présenter :**
  - un panorama *assez large* des algorithmes d'apprentissage artificiel/statistique (réseaux neuronaux [y compris convolutionnels et apprentissage « profond »], SVM, dopage [boosting], forêts aléatoires, HMM, clustering, k-means,...) permettant des analyses et modélisations de données plus puissantes que les simples méthodes statistiques linéaires
  - cadre théorique et méthodologique commun
  - divers types (classification, régression, clustering, ...) et domaines d'applications
- **Faire pratiquer en TPs plusieurs de ces techniques**
- **Travailler sur un mini-projet de machine-learning**
- **Intervenants**
  - Chloé AZENCOTT (CBIO, MINES ParisTech)
  - Fabien MOUTARDE (CAOR, MINES ParisTech)
  - + assistants de TP (Jesus BUJALANCE, Sofiane HORACHE?, ? [doctorants au CAOR])
  - + ?? [doctorants au CBIO])

- Séance 1 (lundi 30 novembre 2020 matin) : *Intro et généralités sur l'apprentissage statistique (rappels et compléments)* [[Chloé AZENCOTT \(CBIO\)](#)]
- Séance 2 (lundi 30 novembre 2020 après-midi) : *Réseaux de neurones à couches (non profonds) = Multi-Layer Perceptron (MLP)* [[Fabien MOUTARDE \(CAOR\)](#)]
- Séance 3 (mardi 1<sup>er</sup> décembre 2020 matin) : *Apprentissage profond (Deep-Learning) et réseaux convolutionnels* [[Fabien MOUTARDE \(CAOR\)](#)]
- Séance 4 (mardi 1<sup>er</sup> décembre 2020 après-midi) : *Apprentissage statistique en Bio-Informatique, SVMs et méthodes à noyau* [[Chloé AZENCOTT \(CBIO\)](#)]
- Séance 5 (mercredi 2 décembre 2020 matin) : *Apprentissage NON- supervisé, clustering* [[Chloé AZENCOTT \(CBIO\)](#)]
- Séance 6 (mercredi 2 décembre 2020 après-midi) : *Arbres de décision (AD) et forêts aléatoires (Random Forests), dopage (boosting)* [[Fabien MOUTARDE \(CAOR\)](#)]
- Séance 7 (jeudi 3 décembre 2020 matin) : *ML pour données séquentielles (e.g. reconnaissance de gestes), HMM, DTW, RNN* [[Fabien MOUTARDE \(CAOR\)](#)]
- Jeudi 3/12/2020 après-midi : *Choix et Démarrage Mini-projets en binômes*
- Vendredi 4/12/2020 matin : *Exam écrit sur table*
- Vendredi 4/12/2020 après-midi : *Mini-projets en binômes (suite)*

## Documents rédigés (faisant fonction de poly)

### Lecture notes:

[Apprentissage Statistique Supervisé, F. Moutarde,](#)  
[article H5010 publié par Techniques de l'Ingénieur](#)  
[\(février 2019\).](#)

### Reference books:

- \* [Introduction au machine learning, C.Azencott, Dunod \(2018\)](#)  
[https://www.dunod.com/sciences-techniques/introduction-  
au-machine-learning-0](https://www.dunod.com/sciences-techniques/introduction-au-machine-learning-0)
- \* [The Elements of Statistical Learning, T. Hastie, R. Tibshirani & J. Friedman](#) (profs à Stanford, où ce livre sert de manuel).  
[[pdf en ligne](#), imprimable sur demande]

# Evaluation : mini-projet + exam écrit

---

## Mini-projet par groupe de 2 élèves

Jeudi 3/12/2020 après-midi = présentation sujets + choix +  
début travail

+ Vendredi 4/12/2020 après-midi = suite du travail en binôme

+ *hors emploi du temps si besoin*

Principe : mettre en œuvre *plusieurs* algorithmes vus en cours sur  
un jeu de données à choisir parmi une liste proposée

Rendu (le 17/12/2020) = un « Jupyter\_notebook » contenant à la fois :

- le code Python écrit et utilisé durant le mini-projet,
- les résultats obtenus (sous forme de tableaux et graphes, etc...)
- une analyse desdits résultats

+ Examen écrit individuel sur table et sans document  
(vendredi 4/12/2020 matin)

## QUELQUES REFERENCES SUR L'APPRENTISSAGE ARTIFICIEL

---

### Introduction au machine learning

C. Azencott, Dunod (2018).

<https://www.dunod.com/sciences-techniques/introduction-au-machine-learning-0>

### The Elements of Statistical Learning (2<sup>nd</sup> edition)

T. Hastier, R. Tibshirani & J. Friedman, Springer, 2009.

<http://statweb.stanford.edu/~tibs/ElemStatLearn/>

### Deep Learning

I. Goodfellow, Y. Bengio & A. Courville, MIT press, 2016.

<http://www.deeplearningbook.org/>

### Pattern recognition and Machine-Learning

C. M. Bishop, Springer, 2006.

### Introduction to Data Mining

P.N. Tan, M. Steinbach & V. Kumar, AddisonWesley, 2006.

### Apprentissage artificiel : concepts et algorithmes

A. Cornuéjols, L. Miclet & Y. Kodratoff, Eyrolles, 2002.