
INTRODUCTION

ANALYSE DES DONNÉES & INTELLIGENCE ARTIFICIELLE/APPRENTISSAGE AUTOMATIQUE
@ L'ACADÉMIE DU NUMERIQUE

LE TSUNAMI DE LA SCIENCE DES DONNÉES

Depuis la venue de l'apprentissage profond (Deep Learning) et des données massives (Big Data), il est devenu essentiel pour les décideur.e.s de se munir **d'équipes dédiées à la science des données.**

L'annonce par le *Harvard Business Review* que de tels spécialistes seraient les « vedettes rock » du 21^{ième} siècle a relancé l'intérêt du publique pour ce domaine.

Où en serons-nous dans 5 ans?

Pourquoi vous retrouvez-vous parmi nous?

L'IDÉAL

Maîtrise en science des données

- 8 cours, 4 semestres, avec devoirs et examens
 - **200** heures en salle de classe + **1000** heures de pratique
 - **cours fondamentaux (5)**
 - concepts fondamentaux (math/stat, programmation, bases de données, communication, préparation des données)
 - introduction à la science des données (survol)
 - analyse des données massives
 - analyse de texte et traitement du langage naturel
 - projet pratique
 - **cours en option (6 cours d'un demi-semestre)**
- point de mire: classification

point de mire: mise en grappe

apprentissage profond I et II

flot de données

analyse multi-media

visualisation avancée

systeme de recommandations

analyse des réseaux

détection d'anomalie

minage d'événements

minage de signatures fréquentes

etc.

Session			I.A./A.A.	Analyse des données
1	2019-01-31	13:00-16:00	Éléments universels de la science des données	
2*	2019-02-07	9:00-12:00	*S.O.	Rudiments de la programmation – Concepts fondamentaux
2	2019-02-07	13:00-16:00	Rudiments de la programmation avec R et Python	
3	2019-02-14	13:00-16:00	Principes statistiques et mathématiques	Architecture des données et de l'information
4	2019-02-21	13:00-16:00	Collecte et traitement des données	
5	2019-02-28	13:00-16:00	Exploration et visualisation des données	
6	2019-03-07	13:00-16:00	Apprentissage statistique; minage par règles d'association	Techniques de base d'analyse des données
7	2019-03-14	13:00-16:00	Labo: Minage par règles d'association	Labo: Techniques de base d'analyse des données
8	2019-03-21	13:00-16:00	Classification et estimation de la valeur	Mesures et paramètres
9	2019-03-28	13:00-16:00	Labo: Classification et estimation de la valeur	Labo: Mesures et paramètres
10	2019-04-04	13:00-16:00	Mise en grappe	Visualisation avancée des données et production de rapports
11	2019-04-11	13:00-16:00	Labo: Mise en grappe	Labo: Visualisation avancée des données et production de rapports
12	2019-04-18	13:00-16:00	Données non structurées et analyse des sentiments	
13	2019-04-25	13:00-16:00	Labo: Données non structurées et analyse des sentiments	
14	2019-05-02	13:00-16:00	Algorithmes avancés; enjeux et défis	Analyse prévisionnelle

OBJECTIFS DU COURS

1. Augmenter votre capacité à discuter de science des données (S.D.), d'intelligence artificielle (I.A.), et d'apprentissage automatique (A.A.)
2. Augmenter vos connaissances fondamentales des domaines S.D., I.A., et A.A.
3. Introduire les concepts fondamentaux de ces domaines
4. Faciliter la compréhension des enjeux et des aspects pratiques de ces domaines
5. Offrir de premières expériences d'analyse pratique
6. Offrir un tremplin vous permettant de guider les 1200 heures qui suivent

OBJECTIFS TANGENTIELS

Les sessions ne portent que **partiellement** sur

- la programmation pour l'analyse des données
- la totalité des méthodes de la S.D., de l'I.A. et de l'A.A.
- les technologies actuelles et les logiciels utilisés (R/Python/etc.) à ces fins

Dans ce cours, la maîtrise de R et de Python **n'est pas essentielle** à l'apprentissage, quoique qu'elle deviendra plus importante dans le contexte pratique. Nous approcherons surtout la programmation par osmose (via les « notebooks »).

APPROCHE PÉDAGOGIQUE

De manière plus générale, les modules d'apprentissage contiennent des **sessions théoriques**, des **exercices pratique** (notebooks R/Python), des **discussions de groupe**, et des **études de cas**.

Les session théoriques sont composées ou de techniques d'analyse, ou d'approches heuristiques, ou encore d'un **mélange** des deux.

Philosophie: il peut s'avérer avantageux de se faire une **idée générale** de l'allure de certaines notions techniques sans nécessairement en connaître tous les détails.

N'hésitez pas à poser des questions.

DÉFIS D'ENSEIGNEMENT

Sans devoirs ou examens... où se situe votre **motivation**?

Vous avez peut-être quitté les bancs d'école depuis longtemps (et avec bonheur...)

Un cours de survol se doit **d'omettre plusieurs détails importants**

Groupes hétérogènes: talents, intérêts, expériences, etc.

Attentes exagérées

Préoccupations **pratiques** vs. préoccupations académiques

“Les premières étapes en éducation *doivent* être parsemées de nombreux petits mensonges (« lies-to-children ») car les explications doivent rester simples au départ. Nous vivons cependant dans un univers d’une grande complexité, et les petits mensonges doivent **être remplacées** (à la longue) par des histoires plus complexes afin d’éviter qu’ils ne deviennent d’authentiques mensonges.”

(T. Pratchett, I. Stewart, J. Cohen, *The Science of Discworld*)

petits mensonges aux enfants: “ce qui est compréhensible”

petits mensonges aux patrons: “ce qui est essentiel à savoir”

petits mensonges aux patients: “ce qui les empêchera de s’inquiéter sans raison”

petits mensonges à nous-mêmes: ...

M Gustave: Je n'en veux pas à Serge. Tout le monde n'a pas la fibre morale. [...] Il n'y est pour rien.

Mustapha Zero: Je ne sais pas, ça dépend.

M Gustave: On peut dire ça de tout. Bien sûr que ça dépend.

Mustapha Zero: Bien sûr que ça dépend.

Mr. Gustave: Oui tu as raison. Bien sûr que ça dépend.

(The Grand Budapest Hotel, 2014)