
INF7100

Initiation à la science des données et à l'intelligence artificielle

Plan de cours

Responsable(s) du cours

Coordination : MEURS, Marie-Jean

Enseignement :

CHARPENTIER, Arthur
PK-5330
poste 8234
charpentier.arthur@uqam.ca

MEURS, Marie-Jean
PK-4935
poste 6139
meurs.marie-jean@uqam.ca
<https://professeurs.uqam.ca/professeur/meurs.marie-jean/>

ROY, Jean-Hugues
J-4185
poste 6102
roy.jean-hugues@uqam.ca
<https://professeurs.uqam.ca/professeur/roy.jean-hugues/>

Description et objectif du cours

L'objectif général est de former une prochaine génération de chercheuses et de chercheurs apte à créer ses propres outils numériques et à comprendre les avantages, mais aussi les limites, de l'informatique en général et de l'«intelligence artificielle» en particulier dans un contexte de recherche dans toutes les disciplines (sciences, sciences humaines et sociales, etc.).

Contenu du cours

Le cours vise l'acquisition de compétences en science des données et IA :

- L'initiation à la programmation au moyen d'un langage informatique d'usage général ;
- La recherche d'information ;
- Le traitement des données massives ;
- La maîtrise de notions fondamentales d'apprentissage automatique.

Le cours comprend des exercices pratiques.

Calendrier

01 - Mardi, 5 mai 2020 - Tous trois

- Description du syllabus
- Test vidéo
- Entente d'évaluation
- Attentes des étudiant.e.s
- Motivations et objectifs du cours
- Connexion à l'environnement : [Carnets Azure](#)

02 - Jeudi, 7 mai 2020 – JH Roy

- Python 1 : introduction générale, types de variables, boucles et conditions

03 - Mardi, 12 mai 2020 – JH Roy

- Python 2 : lire et créer des fichiers

04 - Jeudi, 14 mai 2020 – JH Roy

- Python 3 : interaction avec APIs ; moissonnage de données dans le web

05 - Mardi, 19 mai 2020 – JH Roy

- Python 4 : analyse de données avec pandas

06 - Jeudi, 21 mai 2020 – JH Roy

- Python 5 : antichambre du traitement automatique du langage naturel (TALN) avec NLTK

07 - Mardi, 26 mai 2020 – A Charpentier

- Introduction à la science de données
- Expérimentation vs. observation (courte introduction aux modèles causaux)

08 - Jeudi, 28 mai 2020 – A Charpentier

- Introduction à la visualisation
- Modélisation statistique

09 - Mardi, 2 juin 2020 – A Charpentier

- Corrélation et régression

10 - Jeudi, 4 juin 2020 – A Charpentier

- Algorithmes de classification

11 - Mardi, 9 juin 2020 – MJ Meurs

- Données massives : définition, collecte, traitement
- Applications industrielles, stratégie et régulation

12 - Jeudi, 11 juin 2020 – MJ Meurs

- Recherche d'information et indexation

13 - Mardi, 16 juin 2020 – MJ Meurs

- Introduction à l'apprentissage automatique

14 - Jeudi, 18 juin 2020 – MJ Meurs

- Réglementation, données personnelles

15 - Mardi, 23 juin 2020 - Tous trois

- Présentations des projets

Modalités d'évaluation

Description sommaire	Date de remise / passation sur Moodle	Pondération
Travaux de programmation	Du 07 mai au 25 mai	25%
Travaux science des données	Du 27 mai au 09 juin	25%
Travaux IA	Du 10 juin au 22 juin	25%
Projet de session	Mardi 30 juin avant 23h59	25%

Aucune remise n'est acceptée après la date et l'heure de remise.

L'utilisation de livres et de documentation personnelle est permise pour réaliser les travaux.

Les travaux sont individuels.

Détail des évaluations

- Programmation
 - TP1 - Premiers pas en python (5%)
Remettre votre carnet dans Moodle avant le **11 mai, 9h**.
Inclure le plus de commentaires possibles (ou des blocs de texte en markdown) pour expliquer ce que vous faites.
Donner à votre fichier le nom suivant : *TP1-votre_code_permanent.ipynb*
 - TP2 - Moissonnage (10%)
Remettre votre carnet dans Moodle avant le **18 mai, 9h**.
Inclure le plus de commentaires possibles (ou des blocs de texte en markdown) pour expliquer ce que vous faites.
Donner à votre fichier le nom suivant : *TP2-votre_code_permanent.ipynb*
 - TP3 - Analyse de données dans Pandas (10%)
Remettre votre carnet dans Moodle avant le **25 mai, 9h**.
Inclure le plus de commentaires possibles (ou des blocs de texte en markdown) pour expliquer ce que vous faites.
Donner à votre fichier le nom suivant : *TP3-votre_code_permanent.ipynb*

- Science des données

- TP4 : Moissonnage (10%)
- TP5 : Prédiction (10%) : prévoir une des variables (régression)
- TP6 : Visualisation (5%)
Les modalités de remise seront indiquées pendant la session.

- Données massives et intelligence artificielle

- Synthèse 1 : Données massives - outils et pratiques (10%)
- Synthèse 2 : Indexation - application (5%)
- Synthèse 3 : Apprentissage automatique - étude de cas (10%)
Toutes les remises seront effectuées sur Moodle.

- Projet

Le projet est un travail individuel.

Il pourra porter sur des sujets en lien avec les travaux et recherches de l'étudiant.e.

Le sujet de projet devra être validé par les enseignant.e.s au plus tard le **01 juin**.

Le projet sera présenté oralement lors du dernier cours de la session, le mardi 23 juin.

Le rapport de projet (5 à 10 pages) devra être remis au plus tard le mardi 30 juin.

Répartition de la pondération : La présentation orale vaut pour 5%, le rapport pour 20%.

En cas de plagiat ou de fraude, la sanction peut aller de la note zéro pour le travail ou l'examen, jusqu'à l'exclusion de l'université.

Les règlements concernant le plagiat seront strictement appliqués. Pour plus de renseignements, consultez le site suivant : <http://www.sciences.uqam.ca/etudiants/integrite-academique.html>

Médiagraphie

Gray, J., Bounegru, L. (2019). The Data Journalism Handbook 2 : Towards a Critical Data Practice. En ligne. Récupéré de <https://datajournalismhandbook.org> .

Herzog, D. (2016). Data Literacy. A User's Guide. Thousand Oaks : Sage Publications, Inc. 205 pages.

Cardon, D. (2019). Culture numérique. Paris : Presses de sciences po. 978-2-7246-2367-3

Cairo, A. (2016). *The Truthful Art : Data, Charts, and Maps for Communication*. San Francisco : New Riders. 523 pages.

Wilke, Claus O. *Fundamentals of Data Visualization*. En ligne. <https://serialmentor.com/dataviz/>

Green, B. (2020). *The False Promise of Risk Assessments : Epistemic Reform and the Limits of Fairness*. Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAT*) <https://www.benzevgreen.com/20-fat-risk/>

Green, B. (2019). 'Good' isn't good enough. NeurIPS Joint Workshop on AI for Social Good. <https://www.benzevgreen.com/19-ai4sg/>

L'autorisation de reprendre un examen en cas d'absence est de caractère exceptionnel. Pour obtenir un tel privilège, l'étudiant-e doit avoir des motifs sérieux et bien justifiés.

Il est de la responsabilité de l'étudiant-e de ne pas s'inscrire à des cours qui sont en conflit d'horaire, tant en ce qui concerne les séances de cours ou d'exercices que les examens. **De tels conflits d'horaire ne constituent pas un motif justifiant une demande d'examen de reprise.**

Dans le cas d'une absence pour raison médicale, l'étudiant-e doit joindre un certificat médical original et signé par le médecin décrivant la raison de l'absence à l'examen. Les dates d'invalidité doivent être clairement indiquées sur le certificat. Une vérification de la validité du certificat pourrait être faite. Dans le cas d'une absence pour une raison non médicale, l'étudiant-e doit fournir les documents originaux expliquant et justifiant l'absence à l'examen ; par exemple, lettre de la Cour en cas de participation à un jury, copie du certificat de décès en cas de décès d'un proche, etc. Toute demande incomplète sera refusée. Si la direction du programme d'études de l'étudiant-e constate qu'un étudiant a un comportement récurrent d'absence aux examens, l'étudiant-e peut se voir refuser une reprise d'examen.

L'étudiant-e absent-e lors d'un examen doit, dans les cinq (5) jours ouvrables suivant la date de l'examen, présenter une demande de reprise en utilisant le formulaire prévu, disponible sur le site Web du département à l'adresse suivante : info.uqam.ca/politiques/.

L'étudiant-e doit déposer le formulaire dûment complété au secrétariat de la direction de son programme d'études : PK-3150 pour les programmes de premier cycle, PK-4150 pour les programmes de cycles supérieurs. Pour plus de détails sur la politique d'absence aux examens du Département d'informatique, consultez le site web suivant : info.uqam.ca/politiques.

PLAGIAT Règlement no 18 sur les infractions de nature académique. (extraits)

Tout acte de plagiat, fraude, copiage, tricherie ou falsification de document commis par une étudiante, un étudiant, de même que toute participation à ces actes ou tentative de les commettre, à l'occasion d'un examen ou d'un travail faisant l'objet d'une évaluation ou dans toute autre circonstance, constituent une infraction au sens de ce règlement.

La liste non limitative des infractions est définie comme suit :

- la substitution de personnes ;
- l'utilisation totale ou partielle du texte d'autrui en la faisant passer pour sien ou sans indication de référence ;
- la transmission d'un travail pour fins d'évaluation alors qu'il constitue essentiellement un travail qui a déjà été transmis pour fins d'évaluation académique à l'Université ou dans une autre institution d'enseignement, sauf avec l'accord préalable de l'enseignante, l'enseignant ;
- l'obtention par vol, manoeuvre ou corruption de questions ou de réponses d'examen ou de tout autre document ou matériel non autorisés, ou encore d'une évaluation non méritée ;
- la possession ou l'utilisation, avant ou pendant un examen, de tout document non autorisé ;
- l'utilisation pendant un examen de la copie d'examen d'une autre personne ;
- l'obtention de toute aide non autorisée, qu'elle soit collective ou individuelle ;
- la falsification d'un document, notamment d'un document transmis par l'Université ou d'un document de l'Université transmis ou non à une tierce personne, quelles que soient les circonstances ;
- la falsification de données de recherche dans un travail, notamment une thèse, un mémoire, un mémoire-créditation, un rapport de stage ou un rapport de recherche ;
- Les sanctions reliées à ces infractions sont précisées à l'article 3 du Règlement no 18.

Les règlements concernant le plagiat seront strictement appliqués. Pour plus de renseignements, veuillez consulter les sites suivants : www.sciences.uqam.ca/etudiants/integrite-academique.html et www.bibliotheques.uqam.ca/plagiat/le-plagiat-liens-rapides.

Politique no 16 visant à prévenir et combattre le sexisme et les violences à caractère sexuel**Pour consulter la politique no 16 :**

instances.uqam.ca/wp-content/uploads/sites/47/2018/05/Politique_no_16.pdf

Services offerts :

Pour obtenir de l'aide, faire une divulgation ou une plainte :
Bureau d'intervention et de prévention en matière de harcèlement
514 987-3000, poste 0886

Pour la liste des services offerts en matière de violence sexuelle à l'UQAM et à l'extérieur de l'UQAM : harcelement.uqam.ca

CALACS Trêve pour Elles – point de services UQAM :

514 987-0348

calacs@uqam.ca
trevepourelles.org

Soutien psychologique (Services à la vie étudiante) :

514 987-3185

Local DS-2110

Service de la prévention et de la sécurité : 514 987-3131

Les étudiants qui ont une lettre signée de leur conseillère ou conseiller de l'Accueil et de soutien aux étudiants en situation de handicap (ASESH), dans laquelle il est fait état de leur inscription au ASESH à titre d'étudiant(e) en situation de handicap, sont invités à remettre ce document à leurs professeur(e)s et chargé(e)s de cours dès le début de la session afin que les aménagements dans le respect des exigences académiques soient déterminées de concert avec chacun des professeur(e)s et chargé(e)s de cours. Les étudiants qui ont une déficience et qui ne seraient pas inscrits au ASESH sont priés de se présenter au AB-2300.

Étudiants avant une déficience de type visuelle, auditive, motrice, trouble d'apprentissage, trouble envahissant du développement et trouble de santé mentale :

Les étudiant(e)s qui ont une lettre d'*Attestation des mesures d'aménagements académiques* obtenue auprès d'une conseillère, d'un conseiller de l'ACCUEIL ET SOUTIEN AUX ÉTUDIANTS EN SITUATION DE HANDICAP (ASESH) doivent rencontrer leurs enseignant(e)s au début de la session afin que des mesures d'aménagement en classe ou lors des évaluations puissent être mises en place. Ceux et celles qui ont une déficience ou une incapacité mais qui n'ont pas cette lettre doivent contacter l'ASESH au (514) 987-3148 ou se présenter au AB-2300 le plus tôt possible.