Université de Montréal

Plan de cours

Modifié le 3 septembre 2023

Information générale

Cours	
Titre	Méthodes quantitatives en biologie
Nombre de crédits	3
Sigle	BIO3033-BIO6033
Site StudiUM	Méthodes quantitatives en biologie
Faculté / École / Département	Arts et sciences / Département de sciences biologiques
Trimestre	Automne
Année	2023
Mode d'enseignement	En présentiel
Déroulement du cours	Le cours se déroule au A-3551 (Campus MIL), de 12h30 à 15h30, le mardi à partir du 5 septembre.
Charge de travail hebdomadaire	La charge de travail habituelle se compose des trois heures de cours en présentiel, et sur une base ponctuelle, de lectures et de travaux à faire en dehors du cours.

Enseignant		
Timothée Poisot	**Coordonnées**	Par courriel au timothee.poisot@umontreal.ca Mon bureau est au Campus MIL, B-5415
	Disponibilités	Réservez un créneau horaire ici: https://calendly.com/timotheepoisot/office-hours

Description du cours	
Description simple	Introduction aux bonnes pratiques en sciences des données pour la biologie
Description détaillée	Les objectifs de ce cours sont 1. présentation des bonnes pratiques en gestion, description, et nettoyage des données de recherche 2. présentation des bonnes pratiques en visualisation des données 3. introduction des concepts fondamentaux d'apprentissage automatique 4. survol de la diversité des approches d'apprentissage automatique en biologie 5. illustration des concepts théoriques *via* une étude de cas sur la prédiction des réponses de la biodiversité aux changements climatiques
Place du cours dans le programme	Ce cours est offert en troisième année et aux cycles supérieurs, et est optionnel.

Apprentissages visés

Objectifs généraux	1. Initier au cycle de vie des données de recherche (collecte, nettoyage, vérification, documentation, archivage) 2. Initier aux approches contemporaines en science des données pour les sciences biologiques 3. Développer la capacité de concevoir une étude établie sur des données ouvertes 4. Initier à la programmation avec des données pour générer des prédictions quantitatives
Objectifs d'apprentissage	1. Utiliser des outils de validation des données, notamment *OpenRefine* 2. Choisir un algorithme de *machine learning* en fonction des données et du type de problème 3. Effectuer des *Exploratory Data Analyses* utilisant les approches de visualisation appropriées

	4. Rédiger des documents expliquant le résultat d'une analyse quantitative 5. Programmer des scripts simples pour transformer et analyser des données

Calendrier des séances

5 septembre 2023	**Titre**	Introduction
	Contenus	Présentation du cours et introduction pratique aux concepts principaux Discussion du mode d'évaluation
	Activités	Atelier - c'est quoi, des données? Atelier - c'est quoi, une prédiction?
	Lectures et travaux	Après la séance: installer Open Refine https://openrefine.org
	Évaluation formative	Auto-évaluation pendant les ateliers
	Évaluation	Lecture optionnelle -- plan de gestion des données

12 septembre 2023	**Titre**	Exploration et nettoyage des données
	Contenus	Représentation des données Nettoyage et contrôle de qualité des données
	Activités	Atelier - réconciliation des données Présentation de la plateforme Perusall pour la lecture collaborative
	Lectures et travaux	Lectures et annotations sur Perusall
	Évaluation formative	Évaluation par les pairs pendant l'atelier
	Évaluation	Lectures et annotations (pendant la semaine)

19 septembre 2023	**Titre**	Méthodes de regroupement
	Contenus	Détection de groupes dans les données Mesures de la performance d'un modèle
	Activités	Théorie - la méthode des k-means Démonstration - classification du land use à partir d'images LANDSAT Atelier - installation de Quarto / VS Code
	Lectures et travaux	Lectures et annotations des notes de cours sur Perusall
	Évaluation formative	Rétroaction directe pendant l'atelier
	Évaluation	Lectures et annotations (pendant la semaine)

26 septembre 2023	**Titre**	Régression et descente de gradient
	Contenus	Prédiction de variables quantitatives Gradient et notion d'optimalité Transformation des données
	Activités	Théorie - la descente de gradient Démonstration - propriétés statistiques des réseaux trophiques Atelier - discussion sur la performance du modèle
	Lectures et travaux	Lectures et annotations des notes de cours sur Perusall (incluant les notes annexes sur le data leakage)
	Évaluation formative	Évaluation par les pairs pendant l'atelier
	Évaluation	Lectures et annotations (pendant la semaine) Réplication de l'exemple vu en cours (obl. 6033, facult. 3033)

3 octobre 2023	**Titre**	Validation des modèles
	Contenus	Entrainement, Validation, et Test Validation croisée Fonctions de perte pour les problèmes de régression
	Activités	Théorie - comment valider un modèle prédictif? Démonstration - date de floraison des cerisiers pendant les 1000 dernières années World café - qu'est-ce qu'on (ne) peut (pas) faire avec ce modèle?
	Lectures et travaux	Lectures et annotations des notes de cours sur Perusall
	Évaluation formative	Rétroaction directe pendant le World Café
	Évaluation	Lectures et annotations (pendant la semaine)

10 octobre 2023	**Titre**	La classification
	Contenus	Fonctions de perte pour les problèmes de classification Les modèles de distribution d'espèce: un problème de machine learning?
	Activités	Théorie - la classification binaire Démonstration - distribution spatiale du renne Débat - est-ce qu'on utilise ce modèle?
	Lectures et travaux	Lectures et annotations des notes de cours sur Perusall
	Évaluation formative	Auto-évaluation après le débat
	Évaluation	Lectures et annotations (pendant la semaine) Réplication de l'exemple vu en cours avec des variables différentes (obl. 6033, facult. 3033)

24 octobre 2023	**Titre**	Sélection des variables et dimensionnalité
	Contenus	Méthodes de sélection de variables Réduction de la dimensionnalité Comparaison de modèles
	Activités	Théorie - analyse en composantes principales, data whitening, et stepwise selection Démonstration - comparaison des différentes méthodes Débat - faut-il croire les statistiques ou les intuitions des biologistes?
	Lectures et travaux	Lectures et annotations des notes de cours sur Perusall
	Évaluation formative	Auto-évaluation après le débat
	Évaluation	Lectures et annotations (pendant la semaine)

31 octobre 2023	**Titre**	Courbes d'apprentissage
	Contenus	Ajustement des hyper-paramètres Modèles probabilistes Incertitude et information
	Activités	Théorie - optimisation du seuil dans les modèles probabilistes Démonstration - optimisation du modèle de la distribution du renne World Café - comment parler d'incertitude?
	Lectures et travaux	Lectures et annotations des notes de cours sur Perusall
	Évaluation formative	Rétroaction directe pendant le world café
	Évaluation	Lectures et annotations (pendant la semaine)

7 novembre 2023	**Titre**	Interprétabilité des modèles
	Contenus	Diagrammes de réponses partielles Valeurs de Shapley Approches causales
	Activités	Théorie - comprendre l'effet des variables sur la prédiction Démonstration - qu'est-ce qui explique la distribution du renne? Atelier - comprendre, décrire, prédire, expliquer...?
	Lectures et travaux	Lectures et annotations des notes de cours sur Perusall
	Évaluation formative	Auto-évaluation portant sur la session au complet
	Évaluation	Lectures et annotations (pendant la semaine)

14 novembre 2023	**Titre**	Modèles d'ensemble
	Contenus	Comparaison de modèles Notion de modèle d'ensemble
	Activités	Discussion d'article - Becker et al. (2022) Cette activité va occuper l'ensemble de la séance
	Lectures et travaux	Becker et al. (2022): https://www.thelancet.com/journals/lanmic/article/PIIS2666-5247(21)00245-7/fulltext
	Évaluation formative	Rétroaction pendant le cours
	Évaluation	Aucune évaluation

21 novembre 2023	**Titre**	Projet en équipe - partie 1
	Contenus	Établir une justification pour la création d'un modèle prédictif Préparer un plan d'analyse
	Activités	Théorie - présentation des objectifs du projet Pratique - partie 1 - discussion du jeu de données fourni Pratique - partie 2 - préparation des données
	Évaluation formative	Rétro-action par les pairs pendant la séance
	Évaluation	Plan d'analyse d'une page, rendu en groupe sur StudiUM

28 novembre 2023	**Titre**	Projet en équipe - partie 2
	Contenus	Choix de la stratégie de validation Choix des mesures de performance
	Activités	Pratique - partie 1 - Entraînement du modèle qui servira de référence Pratique - partie 2 - Sélection des variables Pratique - partie 3 - Optimisation des hyper-paramètres
	Évaluation formative	Rétro-action pendant la séance
	Évaluation	Sortie du modèle à remettre par StudiUM

4 décembre 2023	**Titre**	Retours sur le cours et conclusion
	Contenus	Séance de conclusion
	Activités	Think/Pair/Share - comment ces pratiques vont changer notre pratique (personnelle) de la biologie? World Café - comment ces pratiques vont changer notre pratique (collective) de la biologie? Discussion de groupe - modifications a apporter au cours?
	Évaluation formative	Aucune
	Évaluation	Aucune

Attention ! Exceptionnellement, l'enseignant peut apporter des modifications aux dates des évaluations. Le cas échéant, l'enseignant doit obtenir l'appui de la majorité des étudiants de sa classe. Veuillez vous référer à l'article 4.8 du Règlement des études de premier cycle et à l'article 28 du Règlement pédagogique de la Faculté des études supérieures et postdoctorales.

Évaluations

Calendrier des évaluations

11 septembre 2023	**Activité**	Annotation des notes de cours et discussion de groupe sur la plateforme Perusall; cette activité se répète chaque semaine!
	Objectifs d'apprentissage visés	- effectuer un retour sur le matériel vu en cours - démontrer du sens critique en effectuant des commentaires - améliorer sa capacité a échanger autour des approches vues en cours
	Critères d'évaluation	- lecture des notes au complet (1/5) - qualité des commentaires (3/5) - implication dans les discussions (1/5)
	Pondération	Chaque semaine de lecture compte de manière égale dans la note finale. Les deux plus mauvaises semaines ne sont pas comptées dans la note.

23 novembre 2023	**Activité**	Plan d'analyse (1 page) - en groupe
	Objectifs d'apprentissage visés	Faire un lien entre la question biologique et les étapes d'analyse Rédiger un document pour convaincre de la pertinence de l'approche utilisée
	Critères d'évaluation	Grille de notation (sur StudiUM en cours de session)
	Pondération	5% de la note du cours

30 novembre 2023	**Activité**	Sortie du modèle (document Quarto) - en groupe
	Objectifs d'apprentissage visés	Effectuer une analyse complète à partir d'un exemple Discuter des résultats et de l'applicabilité du modèle
	Critères d'évaluation	3 points si le travail est remis 1 point pour la présentation des résultats 1 point pour la discussion des résultats
	Pondération	5% de la note du cours

Attention ! Exceptionnellement, l'enseignant peut apporter des modifications aux dates des évaluations. Le cas échéant, l'enseignant doit obtenir l'appui de la majorité des étudiants de sa classe. Veuillez vous référer à l'article 4.8 du Règlement des études de premier cycle et à l'article 28 du Règlement pédagogique de la Faculté des études supérieures et postdoctorales.

Consignes et règles pour les évaluations

Absence à une évaluation	Un travail de remplacement sera proposé
Dépôts des travaux	En ligne sur la plateforme Perusall, en personne au cours suivant pour les auto-évaluations formatives
Matériel autorisé	Tout matériel autorisé
Qualité de la langue	Travaux écrits acceptés en français, anglais, et espagnol
Seuil de réussite exigé	Ce cours utilise une note seuil de B, qui est garantie si toutes les auto-évaluations sont remises dans le temps

Rappels

Utilisation des technologies en classe	
Enregistrement des cours	L'enregistrement des cours est permis sur tout support qui favorise votre apprentissage; la diffusion de l'enregistrement n'est pas permise.
Prise de notes et activités d'apprentissage avec ordinateurs, tablettes ou téléphones intelligents	Vous pouvez utiliser toute technique de prise de notes qui favorise votre apprentissage.

Ressources

Ressources obligatoires	
Documents	Les notes de cours et les démonstrations seront mises en ligne après chaque séance théorique. Pour les séances qui utiliseraient des diapositives, elles seront partagées avec les annotations.
Ouvrages en réserve à la bibliothèque	N/A

Ressources complémentaires	
Sites Internet	Quarto: https://quarto.org/docs/get-started/ Julia: https://julialang.org VSCode: https://code.visualstudio.com OpenRefine: https://openrefine.org DataViz catalogue: https://datavizcatalogue.com

Soutien à la réussite
De nombreuses activités et ressources sont offertes à l'Université de Montréal pour faire de votre vie étudiante une expérience enrichissante et agréable. La plupart d'entre elles sont gratuites. Explorez les liens ci-dessous pour en savoir plus.
Centre de communication écrite
Centre étudiant de soutien à la réussite
Services des bibliothèques UdeM
Soutien aux étudiants en situation de handicap

Cadres règlementaires et politiques institutionnelles

Règlements et politiques	
Apprenez à connaitre les règlements et les politiques qui encadrent la vie universitaire.	
Règlement des études Que vous soyez étudiant régulier, étudiant libre ou étudiant visiteur, connaitre le règlement qui encadre les études est tout à votre avantage. Consultez-le !	Règlement des études de premier cycle Règlement pédagogique des études supérieures et postdoctorales
Politique-cadre sur l'intégration des étudiants en situation de handicap Renseignez-vous sur les ressources disponibles les mieux adaptées à votre situation auprès du Bureau de soutien aux	Politique-cadre sur l'intégration des étudiants en situation de handicap Demande d'accommodement et responsabilités

étudiants en situation de handicap (BSESH). Le deuxième lien ci-contre présente les accommodements aux examens spécifiques à chaque faculté ou école

Intégrité, fraude et plagiat

Problèmes liés à la gestion du temps, ignorance des droits d'auteurs, crainte de l'échec, désir d'égaliser les chances de réussite des autres – aucune de ces raisons n'est suffisante pour justifier la fraude ou le plagiat. Qu'il soit pratiqué intentionnellement, par insouciance ou par négligence, le plagiat peut entrainer un échec, la suspension, l'exclusion du programme, voire même un renvoi de l'université. Il peut aussi avoir des conséquences directes sur la vie professionnelle future. Plagier ne vaut donc pas la peine !

Le plagiat ne se limite pas à faire passer un texte d'autrui pour sien. Il existe diverses formes de manquement à l'intégrité, de fraude et de plagiat. En voici quelques exemples :

- Dans les travaux : Copier un texte trouvé sur Internet sans le mettre entre guillemets et sans citer sa source ; Soumettre le même travail dans deux cours (autoplagiat) ; Inventer des faits ou des sources d'information ; Obtenir de l'aide non autorisée pour réaliser un travail.
- Durant les évaluations : Utiliser des sources d'information non autorisées ; Obtenir des réponses de façon illicite ; S'identifier faussement comme un étudiant du cours.

Site Intégrité

Les règlements expliqués