

PHY3051: Analyse moderne des données physiques

Analyse des données physiques via l'apprentissage automatique et les techniques modernes de statistique et de données massives

Professeur:

Björn Benneke

Courriel: bjorn.benneke@umontreal.ca

Bureau : CampusMIL B-3007

Auxiliaires d'enseignement:

Thomas Vandal

Courriel: thomas.vandal@umontreal.ca

Bureau : CampusMIL B-3443

Site web du cours:

Ouvert au début du semestre

Aperçu du cours

- L'objectif est de **préparer au mieux les étudiants à leur future carrière dans la recherche scientifique (ou technique)** en leur fournissant tous les outils modernes nécessaires pour comprendre ce que leurs données peuvent leur dire.
- Ce cours porte sur les défis rencontrés dans l'analyse de données physiques (potentiellement de grande dimension) avec des incertitudes expérimentales ou observationnelles.
- Ce cours est pour **les étudiants de 3e années du premier cycle.**
- Le cours enseignera les bases mathématiques, l'utilisation d'outils informatiques modernes et l'interaction avec les mégadonnées.
- Le cours enseignera les bases mathématiques, l'utilisation d'outils informatiques modernes et l'interaction avec les mégadonnées.
- Les étudiants sont supposés connaître **la programmation informatique de base**
 - nous allons développer des compétences supplémentaires avec Python

Objectifs et compétences visés

Après avoir suivi ce cours, les étudiants seront en mesure de :

- Utiliser leur compréhension fondamentale et leurs outils informatiques modernes pour analyser des ensembles de données scientifiques expérimentales et observationnelles.
- Arriver à des conclusions scientifiques en analysant de petits ensembles de données de laboratoire ainsi que des données volumineuses qu'ils rencontreraient régulièrement au cours de leur future carrière universitaire ou professionnelle.
- Comprendre si leurs résultats scientifiques sont solides.
- Utiliser couramment les techniques des statistiques classiques et bayésiennes.
- Rédiger les sections d'analyse des données pour les articles scientifiques.
- Tracer et afficher facilement des ensembles de données complexes et multidimensionnelles.

Manuels de cours (obligatoire)

- 1) Sivia, D.S., Skilling, J., 2006. Data analysis: a Bayesian tutorial. Oxford University Press.
<https://ebookcentral.proquest.com/lib/umontreal-ebooks/detail.action?docID=430582>
https://www.amazon.ca/Data-Analysis-Bayesian-Devinderjit-Sivia/dp/0198568320/ref=sr_1_1?ie=UTF8&qid=1543594837&sr=8-1&keywords=Data+analysis:+a+Bayesian+tutorial
- 2) Gregory, P.C., 2005. Bayesian Logical Data Analysis For The Physical Sciences: A Comparative Approach With Mathematica Support. Cambridge University Press, Cambridge.
<https://www.amazon.ca/s?k=9780521841504&i=stripbooks&linkCode=qs>

Manuels de cours (optionnel)

- 3) Bishop, C.M., 2011. Pattern Recognition and Machine Learning, 1st ed. 2006. Corr. 2nd printing 2011 edition. ed. Springer, New York.
https://www.amazon.ca/gp/product/0387310738/ref=ppx_yo_dt_b_search_asin_title?ie=UTF8&psc=1
- 4) Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. The MIT Press, Cambridge, Massachusetts.
https://www.amazon.ca/Deep-Learning-Ian-Goodfellow/dp/0262035618/ref=asc_df_0262035618/?tag=googleshopc0c-20&linkCode=df0&hvadid=292950359971&hvpos=&hvnetw=g&hvrnd=8433844328291548703&hvpone=&hvptwo=&hvqmt=&hvdev=c&hvdvcm dl=&hvlocint=&hvlocphy=9061029&hvtargid=pla-416263148149&psc=1
- 5) *Kevin P. Murphy, 2023, Probabilistic Machine Learning: Advanced Topics*, The MIT Press, Cambridge, Massachusetts.

Répartition des pourcentage des évaluations

Projet pratique #1	20%
Projet pratique #2	20%
Projet pratique #3	20%
Examen intra	15%
Examen final	25%
Points bonus pour participation active aux discussions que ce soit en classe ou en ligne	

Soumission des projets pratiques:

~~Pendant la période d'enseignement à distance, les devoirs doivent être télécharger sur StudiUM. Tous les codes Python doivent être inclus. S'il vous plaît agraffer toutes les pages ensemble dans une fichier pdf, dans l'ordre des problèmes.~~

Pendant la **période d'enseignement en personne**, les devoirs devraient idéalement être remis sur papier au début de la classe. Tous les codes Python doivent être imprimés et remis avec le reste du travail. S'il vous plaît agraffer toutes les pages ensemble, dans l'ordre des problèmes.

La prolongation du délai de soumission ne peut être demandée que jusqu'à **au moins 5 jours ouvrables avant le délai de soumission initial** et est à la discrétion du professeur. De fortes raisons doivent exister et ne seront exceptées que dans des conditions exceptionnelles.

Forum en ligne (Slack) et questions au professeur

Le professeur mettra en place un forum en ligne "**demandez à vos pairs**". L'objectif de ce forum est que les étudiants s'entraident en posant des questions scientifiques ou python et en y répondant. Les discussions liées à la science et à l'analyse des données sont encouragées. Les contributions au forum seront prises en compte dans la note pour la participation active de chaque étudiant. **On attend de chacun qu'il soit respectueux et poli.**

Il est important de noter que ce forum est là pour poser des questions à vos camarades. **Les questions destinées au professeur doivent être posées dans un courriel directement à bjorn.benneke@umontreal. Veuillez inclure « [PHY3051] » dans la ligne d'objet.** Le professeur répondra à ces questions par courriel dans un délai de 1 à 2 jours ouvrables au mieux de ses capacités. Le professeur peut également choisir de répondre à la question sur StudiUM s'il pense que la question est pertinente pour tous les étudiants. N'hésitez pas à rappeler poliment le professeur si votre question n'a pas reçu de réponse après 2-3 jours ouvrables. Comme les délais pour les projets pratiques sont communiqués plusieurs semaines avant la date limite de soumission, les étudiants devraient poser leurs questions tôt, et idéalement non le dernier jour ouvrable.

~~Pendant les périodes d'enseignement à distance, un forum "random topics" peut être ajouté pour faire du small-talk sur des sujets autres que PHY3051.~~

Calendrier du cours

Semaine	Jour	Sujet	lecture supplémentaire	Projets pratiques
10-Jan	L	Introduction du théorème de Bayes	Sivia & Skilling chapitres 1 - 2.2 et/ou Gregory chapitre 1	
	M	Description formelle de l'incertitude (PDF, CDF) Règles de probabilité		
	V	Exemple d'estimation de paramètres 1 : La mise à jour dans l'analyse bayésienne : le biais d'une pièce de monnaie		
17-Jan	L	Étapes pour faire une analyse bayésienne	Sivia & Skilling chapitres 3.1 - 3.2	
	M	Exemple d'estimation de paramètres 2 : L'amplitude d'un signal en présence d'un bruit de fond (diapos)		
	V	Préparation pour le projet #1 les intégrales pour obtenir la moyenne, la médiane, la matrice de covariance		
24-Jan	L	Comparaison bayésien des modèles	Sivia & Skilling chapitres 4.1 ou/et Gregory chapitre 3	
	M	Comparaison bayésien des modèles		
	V	Travail sur le projet #1		Projet #1
31-Jan	L	Travail sur le projet #1	Sivia & Skilling chapitres 3.6	
	M	Propagation des erreurs		
	V	Pas de TP		
07-Feb	L	Chaîne de Markov Monte Carlo (MCMC): Concepts de base, algorithme (diapos) et démonstration Python de MCMC simple	Gregory chapitre 12	
	M	Chaîne de Markov Monte Carlo (MCMC): Travail pratique en cours: développer le Python Notebook pour résoudre le projet #1		
	V	TP: Préparation pour le projet #2 and batman package		
14-Feb	L	Nested Sampling: Concepts de base	Documentation en ligne de Emcee et du Dynesty	
	M	Nested Sampling: Concepts de base + Dynesty		
	V	TP sur l'évaluation de la convergence des MCMC		
21-Feb	L	Discussion détaillée de la probabilité a-priori	Gregory chapitre 4.1, 4.8	Projet #2
	M	Définition de la fonction de vraisemblance: convolution de bruit dans les données et le modèle		
	V	TP sur l'interprétation des résultats de Nested Sampling et du paquet Dynesty		
28-Feb				

		Semaine de relâche		
07-Mar	L	L'approche fréquentiste vs bayésienne: 1) Description de l'incertitude + 2) vraisemblance maximale et ChiSquared d'une perspective bayésienne	Sivia & Skilling 3.5 et Gregory chapitres 7 -7.2.1	
	M	L'approche fréquentiste vs bayésienne: 1) Traitement des paramètres de nuance; 2) Test de l'hypothèse nulle		
	V	Discussion sur les projets pratiques #3 avec les étudiants de cycle supérieur		
14-Mar	L	Estimation non paramétrique / Processus gaussiens	Murphy 2023 Chapitres 18, Bishop chapitre 5	
	M	Réseaux neuronaux : concepts de base		
	V	TP sur les réseaux neuronaux (concepts de base)		
21-Mar	L	Réseaux de neurones convolutés	Goodfellow chapitre 9 + Documentation en ligne de PyTorch	
	M	PyTorch + Travaux pratiques		
	V	PyTorch + Travaux pratiques		
28-Mar	L	Réseaux neuronaux : concepts de base + les modèles génératifs	Murphy 2023 Chapitres 18	
	M	TP : PyTorch + classification Galaxy		
	V	Les modèles génératifs		Projet #3
04-Apr	L	Les modèles génératifs et inférence basée sur la simulation	Murphy 2023 Chapitres 21-22	
	M	TP sur les modèles génératifs et inférence basée sur la simulation		
	V	Présentations des projets pratiques #3 de PHY6051X		
11-Apr	L	Présentations des projets pratiques #3 de PHY6051X	---	
	M	Présentations des projets pratiques #3 de PHY6051X		
	V	Présentations des projets pratiques #3 de PHY6051X		

Horaire et location d'examen final

XX/XX/20XX, Vendredi 1:30PM - 4:29PM (Examen final) B-2061 MIL Sciences pav. B

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	http://cce.umontreal.ca/
Centre étudiant de soutien à la réussite	http://cesar.umontreal.ca/
Citer ses sources – styles et logiciels (guide)	http://www.bib.umontreal.ca/LGB/
Services du réseau des bibliothèques de l'UdeM	http://www.bib.umontreal.ca/services/default.htm
Soutien aux étudiants en situation de handicap	http://bsesh.umontreal.ca/

Cadres réglementaires et politiques institutionnelles

Règlements et politiques

Apprenez à connaître les règlements et les politiques qui encadrent la vie universitaire d'un étudiant.

Règlement des études

Que vous soyez étudiant régulier, étudiant libre ou visiteur, connaître le règlement qui encadre les études est tout à votre avantage. Consultez-le !

<http://secretariatgeneral.umontreal.ca/documents-officiels/reglements-et-politiques/reglement-des-etudes-de-premier-cycle/>

<http://secretariatgeneral.umontreal.ca/documents-officiels/reglements-et-politiques/reglement-pedagogique-de-la-faculte-des-etudes-superieures-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 étudiants en situation de handicap (BSESH). Le deuxième lien ci-contre présente les accommodements aux examens spécifiques à chaque faculté ou école.

https://secretariatgeneral.umontreal.ca/public/secretariatgeneral/documents/doc_officiels/reglements/administration/adm10_25-politique-cadre_integration_etudiants_situation_handicap.pdf
<http://www.bsesh.umontreal.ca/accommodement/index.htm>

Intégrité, fraude et plagiat

Problèmes liés à la gestion du temps, ignorance des droits d'auteur, 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 entraîner 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 copier-coller ou à regarder la copie d'un collègue. 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 (autoplégat) ; Inventer des faits ou des sources d'information ; Obtenir de l'aide non autorisée pour réaliser un travail.
- Lors des examens : Utiliser des sources d'information non autorisées pendant l'examen ; Regarder les réponses d'une autre personne pendant l'examen ; S'identifier faussement comme un étudiant du cours.

Règlement disciplinaire sur le plagiat ou la fraude <http://www.integrite.umontreal.ca/reglementation/officiels.html>

Site Intégrité <http://integrite.umontreal.ca/>