



PLAN DE COURS

CIV6540 – Méthodes probabilistes d'analyse de données pour ingénieur civil

Hiver 2018 | Département CGM – Polytechnique Montréal

Crédits	3cr (3 - 0 - 6)
Horaire cours	Lundi : 12:45-15:35
Local	A-503
Professeur	James-A. Goulet (bureau B-431.4.6)
Disponibilité	Lundi : 15h35 –16h & Mercredi 15h30 – 17h00

1 Langue d'enseignement

Le cours sera donné en anglais afin de favoriser une familiarisation progressive des étudiants avec la langue de la recherche qui est aujourd'hui l'anglais. Cette familiarisation se veut progressive puisque pendant le cours, les étudiants pourront poser des questions dans les deux langues (anglais et en français) et tous les travaux et examens pourront être faits dans l'une des deux langues, au choix de l'étudiant.

2 Description de l'annuaire

Méthodes d'analyse de données appliquées au domaine du génie civil. Révision des notions de probabilités et statistiques appliquées à l'analyse de données. Méthodes de régression et de classification. Méthodes bayésiennes pour l'estimation des incertitudes épistémiques. Modèles bayésiens linéaires dynamiques pour l'analyse de séquences temporelles de donnée. Méthodes d'échantillonnages Monte-Carlo et d'optimisation convexes pour la calibration de modèles mathématiques et empiriques. Méthodes de propagation des incertitudes à travers les modèles. Identification de décisions optimales à pré-posteriori en fonction de données empiriques.

3 Objectifs

Le cours vise à apprendre à :

- Estimer l'incertitude épistémique liée à un ensemble d'observations empiriques
- Construire des modèles probabilistes à partir de méthodes de régression/classification et d'observations empiriques
- Construire des modèles probabilistes temporels à partir et d'observations empiriques

- Estimer les paramètres de modèles probabilistes à partir de méthodes d'échantillonnage et d'optimisation convexes
- Propager des incertitudes à travers un modèle
- Identifier les décisions optimales dans un contexte incertain
- Utiliser des algorithmes d'analyse de données (Matlab) Appliquer les connaissances acquises des systèmes du génie civil (structures, mines, transport et environnement)

4 Utilité du cours

Ce cours enseignera à l'étudiant comment approcher les problèmes d'analyse de données avec des méthodes probabilistes tirés du domaine de l'apprentissage machine. Le cours est destiné aux ingénieurs civils n'ayant pas un parcours spécialisé en informatique, mathématique ou statistiques. Celui-ci fournira aux étudiants une connaissance des méthodes de base applicable en génie civil.

5 Évaluation

- Travaux pratiques ($5 \times 6\% = 30\%$)
- Contrôle périodique (35%)
- Examen final (35%)

6 Travaux pratiques

L'étudiant exécutera cinq travaux pratiques visant à mettre en pratique la théorie vue en classe. Ces exercices requièrent l'utilisation du logiciel Matlab. Les travaux pratiques seront autocorrigé par les étudiants suivant la procédure expliquée en classe et résumé sur Moodle.

Matière	TP	Contenu
M1–M2	1	Application des concepts de probabilités et statistiques à l'apprentissage machine
M3, M4	2	Estimation Bayésienne des paramètres d'un modèle probabiliste et application de la méthode d'échantillonnage MCMC et de la méthode d'optimisation Newton-Raphson
M5, M6	3	Régression et classification par processus Gaussien
M7	4	Séparation d'une série temporelle en sous-composantes à l'aide du filtre de Kalman
M9, M10	5	Théorie de la décision, Value of information et processus de décision markovien

7 Contenu du cours

Le tableau ci-dessous présente le contenu de chaque module et les lectures correspondantes.

Semaine	Contenu	Lectures
1	Introduction aux défis liés à l'analyse des données en génie civil Révision : Module 1 – Algèbre linéaire, Théorie des ensembles	[1] §2-3 [2] §2 [3] §2.1, 2.2
2	Révision : Module 1 – Théorie des probabilités & fonctions de variables aléatoires	[1] §3 [3] §2.1, 2.2
3	Module 2 : Lois de probabilités appliquées à l'apprentissage machine	[1] §4 [3] §2.4, 2.5
4	Module 3 : Méthodes bayésiennes	[1] §6 [3] §2.2.3, 2.3, 2.4.5, 2.7, 3.3
5	Module 4 : Méthodes d'échantillonnage et d'optimisation	[1] §5,7 [3] §24.1,24.3–24.4
6	Module 5 : Méthodes de régression	[1] §8 [4]
7	Contrôle périodique	
8	Relâche	
9	Module 6 : Méthodes de classification	[1] §9 [5]
10	Module 7 : Analyse de séries temporelles – théorie	[1] §11.1-11.3 [3] §18.1–18.2
11	Module 7 : Analyse de séries temporelles – applications	[1] §11.4-11.5
12	Module 9 : Théorie de la décision	[1] §12
13	Module 10 : Décisions séquentielles et intelligence artificielle	[1] §14-15 [6] §16.1–16.3.2,16.6 [5] §17.1–17.3

8 Programme du cours

L'échéancier général comprenant : les cours, examens, et remise de travaux pratiques (TP) est présenté dans le tableau ci-dessous. À noter que l'échéancier sera être adapté en fonction de la progression du cours.

Semaine	Mois	Date (lundi)	Cours (lundi) module #	Remise électronique des TP (Dimanche, 20h00)	
1	Janvier	8	M0-introduction		
2		15	M1a,b		
3		22	M1b-M2		
4		29	M3		TP1
5	Février	5	M3 & M4		
6		12	M4 & M5		
7		19	M5		TP2
8		26	Contrôle périodique		
9	Mars	5	Semaine de relâche		
10		12	M6		
11		19	M7a	TP3	
12		26	M7b		
13	Avril	2	Paques		
14		9	M9	TP4	
15		16	M10		
16		23	Examen	TP5	

9 Documents & références

Les références sont disponibles sur Moodle

1. Goulet (2018). Probabilistic Machine Learning for Civil Engineers
2. Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep learning. MIT Press
3. Murphy, K. (2012). Machine learning : a probabilistic perspective. The MIT Press
4. Ebden, M. (2008). Gaussian processes for regression : A quick introduction.
5. Ebden, M. (2008). Gaussian processes for classification : A quick introduction.
6. Russell, S. and Norvig, P. (1995). Artificial Intelligence, A modern approach. Prentice-Hall.