

## **ACT 3261 Modélisation prédictive (3 crédits)**

Automne 2022

<b>Professeur :</b>	Philippe Gagnon philippe.gagnon.3@umontreal.ca A.-A. 4241 disponible le jeudi de 14h30 à 16h30
<b>Horaire du cours :</b>	lundi 13h30 à 15h30 au B-3210 Pav. 3200 J.-Brillant jeudi 13h30 à 14h30 au B-4320 Pav. 3200 J.-Brillant
<b>Cours préalable :</b>	ACT 3251 Théorie du risque
<b>Cours concomitant :</b>	STT 3790 Apprentissage statistique

### **Description et objectifs du cours**

L'objectif de ce cours est d'exposer les étudiant(e)s à l'analyse prédictive afin de leur permettre de répondre à des questions comme : qu'est-ce que l'analyse prédictive et en quoi est-ce utile? Quelles sont les techniques d'analyse prédictive couramment utilisées en actuariat? Pour quels types d'analyse ces techniques sont-elles adéquates et quelles sont leurs limites? La majeure partie de la théorie qui sera appliquée et des méthodes qui seront employées a été vue précédemment dans les cours STT 2700 Concepts et méthodes en statistique et STT 2400 Régression linéaire, et de façon concomitante dans STT 3790 Apprentissage statistique. L'exposition à l'analyse prédictive se fera donc par l'intermédiaire de l'application de cette théorie et de ces méthodes, en focussant particulièrement sur la communication des résultats. L'application se fera par le biais d'analyse de jeux de données réels provenant de contextes actuariels, et les résultats seront communiqués sous forme de rapports ou de présentations. Puisque la majeure partie de la théorie qui sera appliquée et des méthodes qui seront employées a été vue précédemment ou de façon concomitante, les séances de cours seront majoritairement utilisées pour présenter et discuter des projets, et pour des présentations de la part des étudiant(e)s, dans le même esprit que des rencontres d'équipe en milieu corporatif. Le cours prépare à l'examen *Predictive Analytics* de la *Society of Actuaries*/Institut canadien des actuaires.

## Contenu du cours

### Contextes d'analyse prédictive et les outils employés

Comprendre dans quels contextes l'analyse prédictive est utile et quels sont les outils permettant de proposer une solution adéquate. Produire une analyse en utilisant le progiciel R à l'aide de RStudio.

### Communication

Dans le cadre d'un rapport ou d'une présentation, justifier de façon simple et claire la solution proposée. Communiquer les résultats de façon simple et claire. Structurer efficacement le rapport ou la présentation.

### Types de données, exploration et résolution de problèmes

Identifier et comprendre les types de données et variables. Explorer et visualiser (à l'aide du *package* R *ggplot2*) les données de façon univariée et bivariée. Comprendre les méthodes de base de traitement des données manquantes et aberrantes.

### Modèles linéaires généralisés

Comprendre les modèles et les hypothèses. Analyser un jeu de données en utilisant ces modèles. Estimer ces modèles et interpréter les résultats. Validation et comparaison de modèles.

### Analyse en composantes principales

Comprendre et appliquer cette technique de réduction de dimension.

### Arbres de décision

Comprendre les arbres de régression et de classification, ainsi que leurs variantes. Analyser un jeu de données à l'aide de ces techniques.

### Techniques de regroupement (*clustering*)

Comprendre et appliquer l'algorithme des *k*-moyennes (*k-means*).

## Références non obligatoires

*An Introduction to Statistical Learning, with Applications in R*, 2ième édition, 2021, par James, G., Witten, D., Hastie, T., Tibshirani, R., New York: Springer (disponible gratuitement à <https://www.statlearning.com/>).

*Data Visualization: A Practical Introduction*, 2018, Healy, K., Princeton University Press (disponible gratuitement à <https://socviz.co/>).

*Healthcare Risk Adjustment and Predictive Modeling*, 2ième édition, 2018, par Duncan, I., ACTEX Learning (chapitre 24, disponible sur Studium).

*R for Everyone*, 2ième édition, 2017, par Lander, J.P., Boston: Addison-Wesley (disponible en version numérique sur le site web de la bibliothèque).

*Regression Modeling with Actuarial and Financial Applications*, 2010, par Frees, E.W., New York: Cambridge (disponible en version numérique sur le site web de la bibliothèque).

## Évaluations

### Cinq projets (50% total, 10% chacun) :

- voir liste avec détails à la fin du document ; les projets seront à faire en équipe.

### Participation (4%) :

- ne vient pas en classe, 0/4;
- vient en classe, 2/4;
- vient en classe et participe, 4/4.

### Un examen (46%) :

- examen final, **lundi le 12 décembre de 13h00 à 16h00 au B-3210 Pav. 3200 J.-Brillant.**

Aucune autre opportunité (devoir, travail supplémentaire, etc.) d'augmenter le nombre de points ne sera allouée. La note finale littérale (en lettre, A+, A, etc.) sera attribuée en fonction de l'atteinte des objectifs spécifiques lors des évaluations.

## Programme d'agrément (règles pour l'année 2022-2023)

Ce cours est agréé en vertu du Programme d'agrément universitaire (PAU) de l'Institut canadien des actuaires (ICA) pour l'année universitaire 2022-2023. Il fait partie des cours nécessaires à réussir pour le crédit du diplôme menant à l'admission pour l'examen synthèse de l'ICA du titre AICA (programme d'agrément avec reconnaissance du diplôme en vigueur dès septembre 2022). Veuillez consulter les pages suivantes pour de plus amples détails :

<https://education.cia-ica.ca/fr/accueil/>  
<https://education.cia-ica.ca/fr/universites/>

En plus des politiques internes en matière de comportements spécifiques à une université, y compris l'inconduite universitaire, les candidat(e)s désirant obtenir des crédits aux examens professionnels sont également assujettis à la Politique relative au Code de conduite et d'éthique des candidats faisant partie du système d'éducation de l'ICA ainsi qu'au Code de conduite et d'éthique pour les candidats au titre d'actuaire dans le système de formation de l'ICA : <https://www.cia-ica.ca/docs/default-source/2016/216106f.pdf> et <http://www.cia-ica.ca/docs/default-source/2016/216107f.pdf>

## Rappels

- La date limite pour modifier un choix de cours coïncide avec la date limite pour abandonner un cours sans frais, soit le 21 septembre.
- La date limite pour l'abandon d'un cours est le 11 novembre.
- Le plagiat : attention, c'est sérieux ! Vous êtes invités à consulter le site internet [www.integrite.umontreal.ca](http://www.integrite.umontreal.ca).

## Calendrier

Septembre							
D	L	M	M	J	V	S	Semaine
				1	2	3	
4	5	6	7	8	9	10	1
11	12	13	14	15	16	17	2
18	19	20	21	22	23	24	3
25	26	27	28	29	30		4
Octobre							
D	L	M	M	J	V	S	Semaine
						1	
2	3	4	5	6	7	8	5
9	10	11	12	13	14	15	6
16	17	18	19	20	21	22	7
23	24	25	26	27	28	29	
30	31						8
Novembre							
D	L	M	M	J	V	S	Semaine
		1	2	3	4	5	8
6	7	8	9	10	11	12	9
13	14	15	16	17	18	19	10
20	21	22	23	24	25	26	11
27	28	29	30				12
Décembre							
D	L	M	M	J	V	S	Semaine
				1	2	3	12
4	5	6	7	8	9	10	13
11	12	13	14	15	16	17	
18	19	20	21	22	23	24	
25	26	27	28	29	30	31	

### Légende

	Examen (final)
	Jours de cours
	Congé
	Semaine de relâche

## Liste de projets avec détails (tentative)

1. Bonnes pratiques en matière de communication des résultats.
  - Date de présentation : 8 septembre.
  - Date de remise : 22 septembre.
2. Types de données, exploration et résolution de problèmes.
  - Date de présentation : 29 septembre.
  - Date de remise : 13 octobre.
3. Régression linéaire : analyse d'un jeu données.
  - Date de présentation : 13 octobre.
  - Date de remise : 3 novembre.
4. Analyse en composantes principales : exploration du même jeu de données.
  - Date de présentation : 3 novembre.
  - Date de remise : 17 novembre.
5. Modèle linéaire généralisé gamma : analyse du même jeu données et comparaison avec la régression linéaire.
  - Date de présentation : 17 novembre.
  - Date de remise : 1 décembre.