



Département d'informatique IFT 870 / BIN 710 – Forage de données

Plan d'activité pédagogique Hiver 2026

Enseignant	Aïda Ouangraoua
Courriel :	aida.ouangraoua@usherbrooke.ca
Local :	D4-1010-7
Téléphone :	+1 819 821-8000 x62014
Disponibilités :	Prendre rendez-vous sur MS Teams

Site web du cours : <https://cobius.usherbrooke.ca/ift870>

Horaire	Exposé magistral :	Mardi	15 h 30 à 16 h 20	salle D3-2034
		Jeudi	10 h 30 à 12 h 20	salle D4-2020

Description officielle de l'activité pédagogique¹

Cibles de formation :	Apprendre les concepts et les techniques de base pour l'entrepôt et le forage de données. Comprendre le rôle du forage dans la réalisation d'un entrepôt de données. Acquérir des habiletés pour développer des systèmes de forage. Appliquer ces concepts et techniques pour la résolution de problèmes.
Contenu :	Entrepôts de données et technologie pour le forage : caractéristiques et architecture, méta-données, modèles multidimensionnels, démarche de construction d'un entrepôt. Processus de forage. Techniques de forage : caractérisation du forage descriptif versus prédictif, prétraitement de données, recherche et extraction des règles d'association, méthodes pour la classification et la prédiction, analyse de faisceaux, évaluation de résultats.
Crédits	3
Organisation	3 heures d'exposé magistral par semaine 6 heures de travail personnel par semaine
Particularités	Aucune

¹<https://www.usherbrooke.ca/admission/fiches-cours/ift870>

1 Présentation

Cette section présente les cibles de formation spécifiques et le contenu détaillé de l'activité pédagogique. Cette section, non modifiable sans l'approbation du comité de programme du Département d'informatique, constitue la version officielle.

1.1 Mise en contexte

Le cours *IFT870 Forage de données (BIN 710 : pour la bio-informatique)* est un cours à option du 2e et du 3e cycle en informatique. Nous y introduisons les principaux aspects du forage de données et nous nous focalisons sur l'utilisation d'un langage tel que Python et d'une bibliothèque (par exemple *scikit-learn*) spécialisés pour le développement d'applications de forage de données. Il ne présuppose aucune connaissance préalable en forage de données ou intelligence artificielle. Les principales méthodes et les aspects pratiques du forage de données seront introduits au fur et à mesure. Des bases en programmation et l'usage de bibliothèques telles que Numpy et Matplotlib sont souhaitables.

1.2 Cibles de formation spécifiques

Ce cours vise des étudiantes et des étudiants qui souhaitent découvrir les principales méthodes de forage de données et approfondir leurs connaissances des aspects pratiques pour le développement d'applications dédiées à des problèmes réels de forage de données. Il ne se focalise pas sur les concepts mathématiques fondamentaux du forage de données, mais plutôt sur les aspects pratiques pour l'utilisation d'algorithmes. À la fin de cette activité pédagogique, l'étudiante ou l'étudiant sera capable :

1. De distinguer et comprendre les algorithmes de forage de données les plus utilisés en pratique ;
2. De maîtriser les étapes de l'implémentation de solutions pour des problèmes de forage de données ;
3. De maîtriser l'utilisation de certains outils informatiques tels que Python, scikit-learn, pandas, matplotlib et les bibliothèques scientifiques de Python pour la mise en oeuvre de solutions.

1.3 Contenu détaillé

Thème	Contenu	Nbr. d'heures	Objectifs	Lectures ¹
1	Introduction : concepts de base, types de données, classification générale des problèmes et algorithmes usuels de forage de données	6	1	Chap. 1 de [2] , [1] et [3]
2	Exploration des données : types de données et d'attributs, descriptions statistiques, mesures de similarité et dissimilarité, visualisation	6	2, 3	Chap. 2 de [2] et [3]
3	Représentation et prétraitement des données : nettoyage, intégration, réduction, transformation et discrétisation des données	6	2, 3	Chap. 3 de [2] et [3] , et 4 de [1]
4	Fonctions prédictives : classification, régression, algorithmes usuels, évaluation, sélection et amélioration de modèles	6	1, 2, 3	Chap. 8,9 de [2] , 2,5 de [1] et 4,5 de [3]
5	Fonctions descriptives : association, corrélations, regroupement, détection d'anomalies, comparaison et évaluation de regroupements	6	1, 2, 3	Chap 10-12 de [2] , 3 de [1] , et 8,10 de [3]
6	Forage de données complexes : données séquentielles, graphes et réseaux, applications du forage de données	6	1, 2	Chap. 13 de [2] , et 7 de [1]

¹ Les lectures indiquées ne sont là qu'à titre indicatif. L'enseignant est libre de choisir un autre document de référence

2 Organisation

Cette section propre à l'approche pédagogique de chaque enseignante ou enseignant présente la méthode pédagogique, le calendrier, le barème et la procédure d'évaluation ainsi que l'échéancier des travaux. Cette section doit être cohérente avec le contenu de la section précédente.

2.1 Méthode pédagogique

Le cours sera donné sous forme d'exposés magistraux. Les exposés magistraux auront pour objectif d'introduire les concepts de base, les algorithmes et les aspects pratiques correspondant avec Python et scikit-learn pour chaque thème, ainsi que faire des exercices. Le cours comporte quatre travaux pratiques. Le but de ces travaux est d'implémenter les concepts et techniques vus en cours.

2.2 Calendrier

Semaine	Commençant le	Thème
1	2026-01-05	1
2	2026-01-12	Exercices et 1
3	2026-01-19	2
4	2026-01-26	2
5	2026-02-02	Exercices et 3
6	2026-02-09	3
7	2026-02-16	Révision et 4
8	2026-02-23	Semaine des examens périodiques
9	2026-03-02	Relâche
10	2026-03-09	4
11	2026-03-16	Exercices et 5
12	2026-03-23	5
13	2026-03-30	6
14	2026-04-06	Révision et 6
15	2026-04-13	Examen
16	2026-04-20	Semaine des examens finals
17	2026-04-27	Semaine des examens finals

2.3 Évaluation

Type de l'évaluation	Pondération	Utilisation des IAG ¹
Devoirs	30 %	Interdite ●
Examen intra	30 %	Interdite ●
Examen final	40 %	Interdite ●

¹ Référez-vous à la page "Balises d'utilisation des outils d'intelligence artificielle générative" à la fin du document.

Les examens se feront à distance. Toute documentation sera autorisée, sous réserve qu'elle soit correctement référencée.

2.3.1 Qualité de la langue et de la présentation

Conformément à l'article 17 du Règlement facultaire d'évaluations des apprentissages² l'enseignante ou l'enseignant peut retourner à l'étudiante ou à l'étudiant tout travail non conforme aux exigences quant à la qualité de la langue et aux normes de présentation.

2.3.2 Plagiat

Le plagiat consiste à utiliser des résultats obtenus par d'autres personnes afin de les faire passer pour sien et dans le dessein de tromper l'enseignante ou l'enseignant. Vous trouverez en annexe un document d'information relatif à l'intégrité intellectuelle qui fait état de l'article 9.4.1 du Règlement des études³. Lors de la correction de tout travail individuel ou de groupe une attention spéciale sera portée au plagiat. Si une preuve de plagiat est attestée, elle sera traitée en conformité, entre autres, avec l'article 9.4.1 du Règlement des études de l'Université de Sherbrooke. L'étudiante ou l'étudiant peut s'exposer à de graves sanctions qui peuvent être soit l'attribution de la note E ou de la note zéro (0) pour un travail, un examen ou une activité évaluée, soit de reprendre un travail, un examen ou une activité pédagogique. Tout travail suspecté de plagiat sera transmis au Secrétaire de la Faculté des sciences. Ceci n'indique pas que vous n'ayez pas le droit de coopérer entre deux équipes, tant que la rédaction finale des documents et la création du programme restent le fait de votre équipe. En cas de doute de plagiat, l'enseignante ou l'enseignant peut demander à l'équipe d'expliquer les notions ou le fonctionnement du code qu'elle ou qu'il considère comme étant plagié. En cas d'incertitude, ne pas hésiter à demander conseil et assistance à l'enseignante ou l'enseignant afin d'éviter toute situation délicate par la suite.

2.4 Échéancier des travaux

Les dates et modalités de remise des travaux seront précisées dans les énoncés.

2.5 Utilisation d'appareils électroniques et du courriel

Selon le règlement complémentaire des études, section 4.2.3⁴, l'utilisation d'ordinateurs, de cellulaires ou de tablettes pendant une prestation est interdite à condition que leur usage soit explicitement permise dans le plan de cours.

Dans ce cours, l'usage de téléphones cellulaires, de tablettes ou d'ordinateurs est autorisé. Cette permission peut être retirée en tout temps si leur usage entraîne des abus.

Tel qu'indiqué dans le règlement universitaire des études, section 4.2.3⁵, toute utilisation d'appareils de captation de la voix ou de l'image exige la permission de la personne enseignante.

Note : Je ne réponds à aucun courriel.

²https://www.usherbrooke.ca/sciences/fileadmin/sites/sciences/documents/Etudiants_actuels/Informations_academiques_et_reglements/2017-10-27_Reglement_facultaire_-_evaluation_des_apprentissages.pdf

³<https://www.usherbrooke.ca/registraire/droits-et-responsabilites/reglement-des-etudes/>

⁴https://www.usherbrooke.ca/sciences/fileadmin/sites/sciences/documents/Etudiants_actuels/Informations_academiques_et_reglements/Sciences_Reglement_complementaire.pdf

⁵<https://www.usherbrooke.ca/registraire/droits-et-responsabilites/reglement-des-etudes/>

3 Matériel nécessaire pour l'activité pédagogique

La matière du cours provient essentiellement des manuels listés dans la section des références. Le Web est également une source de références dans le domaine. Les acétates du cours seront disponibles sur Teams et sur le site Web du cours. Les travaux pratiques seront réalisés dans le langage Python avec les bibliothèques et outils : Numpy, Scipy, Pandas, Matplotlib, et Jupyter Notebook. Les devoirs du cours doivent être remis sous forme de fichiers notebook .ipynb interprétables à l'aide de Jupyter Notebook.

4 Références

- [1] ADREAS C. MÜLLER ET SARAH GUIDO : *Introduction to Machine Learning with Python*. O'Reily Media Inc, 2017.
- [2] JIAWEI HAN, MICHELINE KAMBER ET JIAN PEI : *Data Mining : Concepts and Techniques*. Morgan Kaufmann Publishers Inc, 3rd édition, 2011.
- [3] PANG-NING TAN, MICHAEL STEINBACH ET VIPIN KUMAR : *Introduction to Data Mining*. Addison-Wesley Longman Publishing Co., Inc, 2018.
- [4] PEDREGOSA, FABIAN ET AL. : Scikit-learn : Machine Learning in Python. *Journal of Machine Learning Research*, 12 :2825–2830, 01 2012.

Délits relatifs aux études

Extrait du règlement des études (Règlement 2575-009)

Sont notamment considérés comme un délit relatif aux études les faits suivants :

- a) commettre un plagiat, soit faire passer ou tenter de faire passer pour sien, dans une production évaluée, le travail d'une autre personne, des passages ou idées tirés de l'œuvre d'autrui ou du contenu, de toute forme, généré par un système d'intelligence artificielle (ce qui inclut notamment le fait de ne pas indiquer la source et la référence adéquate);
- b) commettre un autoplagiat, soit soumettre, sans autorisation préalable, une même production, en tout ou en partie, à plus d'une activité pédagogique ou dans une même activité pédagogique (notamment en cas de reprise);
- c) usurper l'identité d'une autre personne ou procéder à une substitution de personne lors d'une production évaluée ou de toute autre prestation obligatoire;
- d) fournir ou obtenir toute forme d'aide non autorisée, qu'elle soit collective ou individuelle (incluant l'assistance provenant d'un système d'intelligence artificielle), pour une production faisant l'objet d'une évaluation;
- e) obtenir par vol ou toute autre manœuvre frauduleuse, posséder ou utiliser du matériel non autorisé de toute forme (incluant le matériel numérique et celui généré par un système d'intelligence artificielle) avant ou pendant une production faisant l'objet d'une évaluation;
- f) copier, contrefaire ou falsifier un document pour l'évaluation d'une activité pédagogique;
- k) posséder ou avoir à sa portée un appareil électronique ou numérique interdit durant une activité d'évaluation;

[...]

Un [guide sur l'intégrité intellectuelle](#) vous est rendu disponible par le service des bibliothèques et des archives de l'Université de Sherbrooke, afin de bien comprendre les différents délits et ainsi éviter d'être aux prises avec un dossier disciplinaire et une ou des sanctions.

Les mesures pouvant être imposées à titre de sanctions disciplinaires sont les suivantes :

- a) la réprimande simple ou sévère consignée au dossier étudiant pour la période fixée par l'autorité disciplinaire ou à défaut, définitivement. En cas de réprimande fixée pour une période déterminée, la décision rendue demeure au dossier de la personne aux seules fins d'attester de l'existence du délit en cas de récidive;
- b) l'obligation de reprendre une production ou une activité pédagogique, dont la note pourra être établie en tenant compte du délit survenu antérieurement;
- c) la diminution de la note ou l'attribution de la note E ou 0;

[...]

Balises d'utilisation des outils d'intelligence artificielle générative

Autorisés ou pas dans les situations d'apprentissage et d'évaluation ?

NIVEAU 0

NIVEAU 1

NIVEAU 2

NIVEAU 3

NIVEAU 4

L'utilisation des outils d'intelligence artificielle générative (IAg) est limitée, voire complètement interdite parce que la personne enseignante considère que l'usage de ces outils nuit au développement de compétences essentielles. Ces compétences peuvent être disciplinaires, comme elles peuvent être d'ordre méthodologique, rédactionnel ou informationnel. Considérant que l'utilisation des IAg requiert un esprit critique, il peut s'agir d'une situation d'apprentissage ou d'évaluation sans IAg qui vise à développer celui-ci.

Dans ces situations, **la personne étudiante produit le travail.**

L'utilisation prononcée des IAg est permise parce que la personne enseignante considère que les personnes étudiantes sont en mesure d'exercer un esprit critique et sont capables de juger de la qualité des contenus produits par les IAg. Ou encore, l'utilisation est encouragée parce que la situation d'apprentissage ou d'évaluation proposée contribue à développer leur esprit critique.

Dans ces situations, l'IAg produit le travail préliminaire, alors que **la personne étudiante s'assure de sa qualité en l'améliorant.**



Utilisation interdite

Le **NIVEAU 0** signifie que l'**utilisation est interdite**.

Ceci signifie que si la personne enseignante a un motif de croire qu'il y a eu l'utilisation d'une IAg dans une situation d'évaluation, elle doit dénoncer les faits auprès de la personne responsable des dossiers disciplinaires universitaires. Il s'agit d'un délit relatif aux études tel que stipulé dans le [Règlement des études](#).



Utilisation limitée

Le **NIVEAU 1 D'UTILISATION** signifie que l'**utilisation est autorisée uniquement pour assister l'apprentissage dans le domaine disciplinaire ou des langues**.

Dans ce contexte, la personne étudiante **est tenue de déclarer l'utilisation qu'elle en a faite** selon les consignes fournies par la personne enseignante sans quoi l'utilisation peut être considérée comme un délit. Par exemple :

Domaine disciplinaire :

- S'inspirer
- Générer des idées
- Explorer un sujet pour mieux le comprendre
- Générer du matériel pour apprendre

Domaine des langues :

- Identifier ses erreurs et se les faire expliquer
- Reformuler un texte
- Générer un plan pour aider à structurer un texte
- Traduire un texte



Utilisation guidée

Le **NIVEAU 2 D'UTILISATION** signifie que l'**utilisation est autorisée pour améliorer un travail produit par la personne étudiante**.

Dans ce contexte, la personne étudiante **est tenue de déclarer l'utilisation qu'elle en a faite** selon les consignes fournies par la personne enseignante sans quoi l'utilisation est considérée comme un délit. Par exemple :

- Analyser des contenus
- Obtenir une rétroaction
- Évaluer la qualité de son travail à partir de critères
- Demander à être confronté relativement à ses idées, à sa démarche
- Diriger les processus de résolution de problèmes



Utilisation balisée

Le **NIVEAU 3 D'UTILISATION** signifie que l'**utilisation est autorisée pour produire un travail qui sera amélioré**.

Dans ce contexte, la personne étudiante **est tenue de citer selon les normes¹ le contenu généré par l'IAg ou de déclarer l'utilisation qu'elle en a faite** selon les consignes fournies par la personne enseignante sans quoi l'utilisation est considérée comme un délit. Par exemple :

- Résumer ou rédiger des parties d'un texte
- Générer un texte ou un modèle d'une production et l'adapter
- Réaliser des calculs mathématiques
- Produire du code informatique
- Résoudre des problèmes complexes
- Répondre à une question
- Générer des images, ou autres contenus multimédias



Utilisation libre

Le **NIVEAU 4 D'UTILISATION** signifie qu'**aucune restriction spécifique n'est imposée**.

Dans ce contexte, la personne étudiante **est tenue de citer selon les normes¹ le contenu généré par l'IAg ou de déclarer l'utilisation qu'elle en a faite** selon les consignes fournies par la personne enseignante sans quoi l'utilisation est considérée comme un délit.

Ce niveau inclut tout ce qui précède, de l'exploration à la production, ainsi que toute autre tâche particulière jugée complexe.

À considérer avant l'utilisation d'outils d'intelligence artificielles génératives

Si, en tant que personne étudiante envisagez d'utiliser un outil d'intelligence artificielle générative (IAG) lorsque l'évaluation autorise les niveaux 1 à 4 d'utilisation mentionnés précédemment.

Dans ce cas, gardez à l'esprit les éléments clés suivants.

- Vous assumez la responsabilité de tout le contenu produit, avec ou sans IAG, et intégré à votre production.
- Les produits des outils d'IAG peuvent très souvent comporter **des erreurs ou des faussetés** (hallucinations) : on doit donc impérativement valider tout contenu généré par ces outils.
- Dans l'état actuel de la Loi sur le droit d'auteur du Canada, les **productions faites par l'IAG sont du domaine public**, puisque les outils d'IAG ne sont pas reconnus comme des auteurs au sens de la Loi et que les contenus générés ne répondent pas aux critères d'une œuvre protégée, notamment aux critères d'originalité.
- L'entreprise qui fournit le service pourrait émettre certaines exigences dans ses conditions d'utilisation. Comme l'algorithme et le code informatique appartiennent à l'entreprise qui les a développés, nous devons tenir compte de ces conditions. Celles-ci pourraient également fournir des précisions relatives à la **réutilisation des données soumises (confidentialité)**.

Comment déclarer l'utilisation d'outils d'intelligence artificielle générative

Dans l'esprit d'une conduite intègre et responsable, vous devez TOUJOURS mentionner de façon explicite toute utilisation de l'intelligence artificielle, conformément au Règlement des études (9.4.1 Délits relatifs aux études). De plus, à des fins pédagogiques, il est recommandé de toujours intégrer à la production les requêtes, de même que les réponses intégrales générées par les outils d'IAG. Celles-ci pourront être intégrées directement dans le corps du texte ou en note de bas de page. Les réponses longues pourraient être insérées en annexe de votre document ou dans des documents supplémentaires, selon les directives de la personne enseignante.

L'utilisation de ces deux documents s'avèrera utile, ils se trouvent sous licence libre, donc vous pouvez utiliser les tableaux et les adapter selon votre besoin:

1. Modèle de citation : Ce formulaire, à remplir par l'enseignant, donne un exemple aux étudiants de citation de l'IAG dans la réalisation d'un travail évalué ou non.
2. Déclaration d'usage : Ce formulaire, à remplir par les étudiants, doit être remis avec une réalisation afin de déclarer l'usage de l'IAG dans la réalisation, qu'elle soit évaluée ou non.

Référence

La Faculté des sciences tient à remercier le SSF pour la production des documents.

- Cabana, M. et Côté, J.-A. (2024). Balises d'utilisation des outils d'intelligence artificielle générative. Service de soutien à la formation, Université de Sherbrooke. Sous licence [CC BY 4.0](#).
- Cabana, M. et Beaudet, M. (2024). Directives de déclaration de l'utilisation de l'intelligence artificielle générative dans une production étudiante. Service de soutien à la formation, Université de Sherbrooke. Sous licence [CC BY 4.0](#).
- Cabana, M. (2024). Formulaire de déclaration de l'utilisation de l'intelligence artificielle générative dans une production étudiante. Service de soutien à la formation, Université de Sherbrooke. Sous licence [CC BY 4.0](#).