

Sigle : INF6243 Gr.01
Titre : Techniques d'apprentissage
Session : Hiver 2025 Horaire et local
Professeur : Tajeuna, Etienne Gael

1. Description du cours paraissant à l'annuaire :

Objectifs

Permettre aux étudiants de maîtriser les concepts fondamentaux de l'apprentissage automatique et d'appliquer ces notions à des problèmes concrets. Leur faire acquérir des connaissances sur les techniques d'apprentissage supervisé et non supervisé, les techniques d'apprentissage pour les données textuelles, les algorithmes de classement des pages Web.

Contenu

Concepts d'apprentissage supervisé : classification et régression, frontière de décision et fonctions discriminantes; Arbres de décision et techniques de traitement du sur-apprentissage (overfitting); Apprentissage par ensemble : (bagging), (boosting) et forêt d'arbres; Machine à noyaux : dimension VC et machines à supports vectorielles; Apprentissage non supervisé : (clustering), les mélanges de loi de distribution statistique, carte de Kohonen et algorithme SOFM; Apprentissage de données multidimensionnelles : techniques de réduction de la dimension, classification non supervisée dans les sous-espaces de dimension (subspace clustering); Fouille de données textuelles : modèle TF-IDF et analyse sémantique latente; Prospection du Web : algorithmes HITS et PageRank.

Descriptif – Annuaire

2. Objectifs spécifiques du cours :

Au terme de cette activité, l'étudiant(e) aura acquis des connaissances sûres :

- Les principes de l'apprentissage automatique;
- Les principes de l'apprentissage supervisé et non supervisé;
- Les méthodologies de régression, de classification, ainsi que leur validation;
- Les techniques de groupement de données et de réduction de dimensions;
- Les applications de l'apprentissage automatique.

3. Stratégies pédagogiques :

Les formules pédagogiques suivantes seront utilisées :

- Cours magistraux (3 heures par semaine) en mode présentiel.
- Ateliers pratiques avec Python (6 séances de 2 heures chacune) en mode présentiel.

4. Heures de disponibilité ou modalités pour rendez-vous :

Bureau : A2222
Courriel : etiennegael.tajeuna@uqo.ca

5. Plan détaillé du cours sur 15 semaines :

Semaine	Thèmes	Dates
1	Partie I - Introduction Introduction à l'apprentissage automatique <ul style="list-style-type: none"> • Données massives (Big data) et apprentissage automatique 	17 Janv. 2025

	<ul style="list-style-type: none"> • Principes et applications de l'apprentissage automatique • Rappel sur le calcul des probabilités et statistique 	
2	<p>Partie I (suite et fin)</p> <p>Introduction à la théorie de l'apprentissage</p> <ul style="list-style-type: none"> • Types d'apprentissage automatique • Minimisation du risque empirique d'apprentissage • Borne sur l'erreur de généralisation d'apprentissage <p>Atelier 1 : Introduction à python</p>	24 Janv. 2025
3	<p>Partie II – Ingénierie des caractéristiques</p> <ul style="list-style-type: none"> • Sélection de caractéristiques • Analyse en composantes principales/indépendantes • TSNE / UMAP • Les auto-encodeurs 	31 Janv. 2025
4	<p>Partie III - Apprentissage supervisé</p> <p>Modèles linéaires pour la régression/classification</p> <ul style="list-style-type: none"> • Régression et classification linéaires • Sous-apprentissage et sur-apprentissage • Validation de l'apprentissage automatique <p>Atelier 2 : Prétraitement et visualisation des données</p> <p>Projet P1</p>	07 Fév. 2025
5	<p>Partie III (suite)</p> <p>Classification supervisée I</p> <ul style="list-style-type: none"> • Algorithme des K plus proches voisins (KPPV) • Classification par arbres de décision • Classification de Bayes naïf • Classification par régression logistique • Machine à vecteurs de support (SVM) <p>Atelier 3 : Apprentissage supervisé : Régression</p>	14 Fév. 2025
6	<p>Partie III (suite)</p> <p>Classification supervisée II</p> <ul style="list-style-type: none"> • Réseaux de neurones artificiels • Apprentissage par propagation en arrière • Algorithme de descente du gradient et mini-batch • Apprentissage profond <p>Atelier 4 : Apprentissage supervisé : Classification</p>	21 Fév. 2025

7	Préparation examen de mi-session Projet P2	28 Fév. 2025
8	Semaine d'étude	3 au 7 mars 2025
9	Examen de mi-session	14 mars 2025
10	Partie IV – Apprentissage non-supervisé Algorithmes de regroupement de données <ul style="list-style-type: none"> • Regroupement hiérarchique de données • Méthode spectrale • Algorithme des K-moyennes • Modèle de mélange Gaussiens • Nombre optimal de groupes • Regroupement par projection Atelier 5 : Regroupement de données	21 mars 2025
11	Partie V – Apprentissage avancé Concepts avancés de régression/classification <ul style="list-style-type: none"> • Algorithmes de fonctions d'ordonnancement (ranking) • Algorithmes ensemblistes (Boosting vs. Bagging) • Algorithmes semi-supervisés • Apprentissage des représentations Atelier 6 : Modèles avancés en classification Projet P3	28 mars 2025
12	Partie V (suite) Analyse et classification de données textuelles et Web <ul style="list-style-type: none"> • Analyse de documents textuels • Modèle TF-IDF et classification vectorielles • Plongement lexical (<i>word embedding</i>) • Algorithmes basés sur l'apprentissage profond 	04 avril 2025
13	Partie V (suite et fin) Analyse et classification de données visuelles <ul style="list-style-type: none"> • Introduction à la vision artificielle • Extraction de caractéristiques visuelles • Classification et reconnaissance visuelles 	11 avril 2025
14	Vendredi saint	18 avril 2025
15	Examen final	25 avril 2025

6. Évaluation du cours :

Dans ce cours, l'évaluation des étudiant(e)s se fera à travers des ateliers, des projets, ainsi que des examens de mi-session et final. La répartition de la note finale sera la suivante :

- **Livrables atelier (05) : $5 \times 2 = 10$ %**
- **Projets (03) : $10 + 10 + 15 = 35$ %**
- **Examen de mi-session : 20 %**
- **Examen final : 35 %**

De manière détaillée, l'évaluation des ateliers et projets sera élaborée comme suit :

- **Livrables ateliers :**

Les séances d'ateliers visent à appliquer les concepts abordés en cours magistraux. Les étudiants travailleront exclusivement avec le langage Python pour implémenter et tester des algorithmes d'apprentissage automatique. À l'issue de chaque séance, un code fonctionnel devra être soumis dans un délai maximum de 24 heures. Chaque livrable sera noté et représentera **2 %** de la note finale.

- **Projets :**

Au cours du trimestre, les étudiants réaliseront trois projets, dont le niveau de difficulté sera adapté à la progression des cours magistraux. Les deux premiers projets seront évalués à 10 % chacun, représentant ensemble 20 % de la note finale, tandis que le troisième projet, accompagné d'une présentation de 20 minutes, comptera pour 15 %. Les projets pourront être effectués individuellement ou en équipes de deux ou trois étudiants si le nombre d'inscrits dépasse 12. Les consignes spécifiques pour chaque projet seront fournies lors de leur lancement, mais tous les rapports devront respecter le format suivant : page de garde, marges de 1 pouce, interligne de 1,5 et police de taille 12. Le libellé du projet **PX**(= 1, 2 ou 3) par l'équipe **N** doit être **INF6243-PX-EquipeN**.

La note de passage requise est de 54%.

Tout retard dans la remise d'un travail entraîne une pénalité de **15 % par jour** sur la note attribuée à ce travail, **jusqu'au maximum d'une semaine**. La qualité du français sera considérée lors de la correction des travaux.

7. Politiques départementales et institutionnelles :

- Politique du département d'informatique et d'ingénierie relative à la tenue des examens
- Note sur le plagiat et sur la fraude
- Politique relative à la qualité de l'expression française écrite chez les étudiants et les étudiantes de premier cycle à l'UQO
- Absence aux examens : cadre de gestion, demande de reprise d'examen (formulaire)

Tolérance **ZÉRO** en matière de violence à caractère sexuel.

Le Bureau d'intervention et de prévention en matière de harcèlement (BIPH) a pour mission d'accueillir, soutenir et guider toute personne vivant une situation de harcèlement, de discrimination ou de violence à caractère sexuel. Le BIPH oriente ses actions afin de prévenir les violences à caractère sexuel pour que nous puissions étudier, travailler et s'épanouir dans un milieu sain et sécuritaire.

Vous vivez ou êtes une personne témoin d'une situation de violence à caractère sexuel ? Vous êtes une personne membre de la communauté étudiante ou une personne membre du personnel, autant à Gatineau qu'à Ripon et St-Jérôme, l'équipe du BIPH est là pour vous, sans jugement et en toute confidentialité.

Ensemble, participons à une culture de respect.

Pour de plus amples renseignements consultez [UQO.ca/biph](https://uqo.ca/biph) ou écrivez-nous au Biph@uqo.ca

8. Principales références :

1. E. Alpayedin. Introduction to Machine Learning, fourth edition. MIT Press, 4th Edition, 2020.
2. Yuxi Hayden Liu. Python Machine Learning by Example. Packt, 2020.
3. MASSIH-REZA AMINI. Apprentissage machine : de la théorie à la pratique. EYROLLES, 2015.
4. C. Bishop. Pattern Recognition and Machine Learning. Springer 2006.
5. C.D. Manning, P. Raghavan et H. Schütze, Introduction to Information Retrieval, Cambridge University Press. 2008.

9. Page Web du cours :

<https://moodle.uqo.ca>