

Sigle : INF6193 Gr. 01

Titre : Intelligence d'affaires

Session : Automne 2022 Horaire et local

Professeur : Missaoui, Rokia

1. Description du cours paraissant à l'annuaire :

Objectifs

Permettre aux étudiants de maîtriser les fondements, concepts et problèmes reliés à l'intelligence d'affaires (veille économique ou Business intelligence) qui inclut trois facettes : la visualisation de l'information et de la connaissance, la fouille de données et l'entreposage de données.

Contenu

Fouille de données : étapes de découverte de connaissances (prétraitement, fouille de données et interprétation des résultats), techniques de classification (arbres de décision, etc.), techniques de regroupement (treillis de concepts, classification hiérarchique), règles d'association, motifs séquentiels, cas aberrants et fouille de données complexes (Web, texte, graphe). Entreposage de données : étapes de construction d'un entrepôt de données (acquisition, stockage, traitement et accès), modélisation multidimensionnelle des données, techniques OLAP, types d'architectures des entrepôts de données, optimisation des performances et matérialisation de cubes de données. Visualisation de l'information et de la connaissance. Intégration des deux technologies de fouille de données et entreposage de données. Applications et outils de la veille économique.

Descriptif – Annuaire

2. Objectifs spécifiques du cours :

Ce cours est destiné aux étudiant(e)s ayant suivi au préalable un ou des cours de bases de données au niveau baccalauréat et/ou qui ont une expérience pertinente en conception et utilisation des bases de données ou en apprentissage machine (*machine learning*). Ce cours vise à approfondir les connaissances de l'étudiant(e) sur des thèmes reliés à l'intelligence d'affaires.

Ce cours a comme objectifs plus spécifiques :

- d'initier l'étudiant(e) aux fondements, concepts et problèmes reliés à la veille économique;
- de permettre à l'étudiant(e) d'approfondir ses connaissances sur les techniques de fouille et d'entreposage de données à travers les cours magistraux, des synthèses d'articles, des travaux pratiques (TP) et le projet de session;
- d'initier l'étudiant(e) à un travail de recherche similaire à la préparation d'un mémoire de maîtrise par la production de références bibliographiques, la préparation de synthèses d'articles et d'une présentations orale, et la rédaction d'un rapport de session.

3. Stratégies pédagogiques :

Il s'agit d'un cours magistral en non-présentiel avec la plate-forme Zoom et des travaux d'équipe **sauf le jour de l'examen**.

Les étudiants qui s'inscrivent à ce cours doivent s'assurer qu'ils ont : un ordinateur (avec un système d'exploitation Windows); une connexion Internet; une webcam; un microphone; la suite Office 365 (les étudiants ont un accès gratuit à la suite Office 365 : <https://uqo.ca/sti/outils-numeriques>).

Guide d'utilisation de Zoom à l'intention des étudiants

Site pour soutien de réussite en mode non-présentiel : uqo.ca/etudier-non-presentiel.

Nota : On pourra vous demander d'installer quelques logiciels selon les activités, mais il s'agira de logiciels gratuits.

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

Période de **consultation** sur **rendez-vous** : les mercredis de 16 h à 17 h 30

5. Plan détaillé du cours sur 15 semaines :

Semaine	Thèmes	Dates
1	Présentation du cours et introduction à l'intelligence d'affaires	8 sept. 2022
2	Techniques de prétraitement de données : épuration, intégration et transformation	15 sept. 2022
3	Découverte de connaissances : raison d'être et étapes <i>Formation des équipes</i>	22 sept. 2022
4	Techniques de fouille de données et applications TP en prétraitement de données : 10 points	29 sept. 2022
5	Classification et prédiction : arbres de décision, réseaux Bayésiens, régression et mesures de précision	6 oct. 2022
6	Semaine d'études <i>Liste des projets de session</i>	13 oct. 2022
7	Regroupement (<i>clustering</i>) : classification hiérarchique, <i>K-means</i> , analyse formelle de concepts	20 oct. 2022
8	Règles d'association et mesures de qualité TP en fouille de données : 10 points	27 oct. 2022
9	Étapes de construction d'un entrepôt de données Ébauche du projet de session	03 nov. 2022
10	Modélisation multidimensionnelle et stratégies de conception	10 nov. 2022
11	Techniques OLAP	17 nov. 2022
12	Architectures des ED et optimisation des performances TP en entreposage de données : 10 points	24 nov. 2022
13	Présentations orales : 15 points	1 ^{er} déc. 2022
14	Présentations orales (fin)	8 déc. 2022
15	Examen final en présentiel : 30 points Rapport final : 25 points	15 déc. 2022

6. Évaluation du cours :

Les normes de présentation de travaux (p. ex. page de garde, marge d'un pouce, interligne à 1.5, taille des caractères de 12 points) doivent être absolument respectées. Tout travail non conforme aux consignes se verra attribuer une pénalité de 10 % de la note obtenue.

Tout travail soumis nécessairement via Moodle en format PDF aura pour nom **INF6193-TPX-EquipeY** ou **INF6193-TPX-Nom**. Aucun travail remis par courriel ne sera pris en considération.

Liste des travaux :

1. TP en prétraitement de données pour le **29 septembre : 10 points**
2. TP en fouille de données (travail en équipe) pour le **27 octobre : 10 points**. Il s'agit d'utiliser un outil pour expérimenter un ensemble de méthodes de fouille de données,
3. TP en entreposage de données (travail en équipe) pour le **24 novembre : 10 points**
Il s'agit de mettre en pratique des techniques de modélisation multidimensionnelle des entrepôts de données.
4. Examen final le **15 décembre : 30 points**
L'examen final porte sur toute la matière étudiée dans le cours et peut être à livre ouvert ou fermé. Il sera de préférence en mode présentiel.
5. Projet de session (travail en équipe) : **40 points**
Les étudiant(e)s sont invité(e)s à proposer et à réaliser un projet d'intelligence d'affaires en équipe de deux membres. Chaque équipe est responsable d'identifier le sujet de son projet et de valider ce sujet auprès du professeur.
La liste des sujets de projets sera finalisée à *la sixième semaine* de la session.

Le projet de session peut être soit théorique (synthèse d'articles), soit pratique (p. ex. conception et implantation d'algorithmes en intelligence d'affaires). Il comprend les trois composantes suivantes :

- Ébauche du projet de session pour le **3 novembre**
- Présentations orales pendant les séances du **1^{er} et du 8 décembre : 15 points**
- Rapport final de 15 pages au maximum pour le 15 décembre : **25 points**

Voici quelques suggestions de sujets de projets de session :

- Intelligence d'affaires et données massives (*big data*)
- Techniques avancées de l'OLAP
- *Power BI* de Microsoft
- Fouille du Web
- Apprentissage profond (*Deep learning*)
- Analyse et fouille de réseaux sociaux
- Techniques de détection de fraudes
- Optimisation des requêtes dans les entrepôts de données
- Entrepôts de données (ED) temporels, spatiaux, distribués ou en temps réel
- Nouvelles technologies en BI
- Science de données

Les consignes de réalisation des divers travaux feront l'objet d'un document à part. Tout retard dans la remise d'un travail entraîne systématiquement une pénalité de 5 % par jour sur la note attribuée à ce travail.

À la discrétion du professeur, une pénalité maximale de 10 % de la note globale pourrait affecter la note de session si le professeur observe des absences fréquentes ou une passivité injustifiée et évidente au cours. 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)

La communauté universitaire s'engage à lutter contre les inconduites, le harcèlement et les violences à caractère sexuel. Dénonçons toute forme de violence.

Ensemble, accomplissons un pas de plus en complétant la formation obligatoire en ligne : "La banalisation des violences à caractère sexuel".

uqo.ca/bimi/formation-obligatoire

Pour de plus amples renseignements consultez :

bimi@uqo.ca



8. Principales références :

Références recommandées pour le cours

Jiawei Han, Micheline Kamber, Jian Pei. *Data Mining Concepts and Techniques*, 3rd Edition, Morgan Kaufmann, 2011.

Ralph Kimball, Margy Ross. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 2013.

Alejandro Vaisman & Estaban Zimanyi. *Data Warehouse Systems – Design and Implementation*, Springer, 2014. Une version électronique du livre existe sur le site Web de la bibliothèque de l'UQO.

Rick Sherman. *Business Intelligence Guidebook: From Data Integration to Analytics*, 1st Edition, Morgan Kaufmann, 2015.

Références additionnelles

Carlo Batini & Monica Scannapieco. *Data and Information Quality – Dimensions, Principles and Techniques*, Springer, 2016.

Matteo Golfarelli & Stefano Rizzi. *Data Warehouse Design: Modern Principles and Methodologies*, McGraw-Hill, 2009.

W.H. Inmon & Dan Linstedt. *Data Architecture: A Primer for the Data Scientist: Big Data, Data Warehouse and Data Vault*, Morgan Kaufmann, 2015.

Krish Krishnan. *Data Warehousing in the Age of Big Data – The Morgan Kaufmann Series on Business Intelligence*, 2013.

Witten, H. & Frank, E., Hall, M.A. & Pal, C.J. *Data Mining: Practical Machine Learning Tools and Techniques*, 4th Edition, Morgan Kaufmann, 2017.

Références en ligne

- Site le plus complet sur la découverte de connaissances : <https://www.kdnuggets.com/>
- Site de Prof. R. Godin de l'UQAM : http://www.labunix.uqam.ca/~godin_r/SiteWebGodin2012/index.html
- Site de Prof. Jiawei Han : <http://hanj.cs.illinois.edu/>
- Services d'analyse de Microsoft SQL Server (*SQL Server Analysis Services*)
<https://docs.microsoft.com/fr-fr/analysis-services/analysis-services-overview?view=sql-analysis-services-2022>
- Site de la plateforme *RapidMiner de science de données* : <https://rapidminer.com/>
- Groupe d'intérêt SIGMOD de l'ACM : <http://www.sigmod.org/>
- *Google Scholar* : <http://scholar.google.ca/>
- *The DBLP Computer Science Bibliography* : <http://www.informatik.uni-trier.de/~ley/db/>
- *ACM Digital Library* : <https://dl.acm.org/> (*ACM Digital Library*)
- Grand dictionnaire terminologique de l'Office québécois de la langue française
<http://gdt.oqlf.gouv.qc.ca/index.aspx>
- Dictionnaire *Linguee* anglais-français : <https://www.linguee.fr/>
- Termium Plus - La banque de données terminologiques et linguistiques du gouvernement du Canada
<https://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-fra.html?lang=fra>

9. Notes du cours site Web du cours :

Disponibles sur Moodle : <https://moodle.uqo.ca>