

5.00 crédits

30.0 h

Q1

Enseignants	Fouss François ;
Langue d'enseignement	Français
Lieu du cours	Mons
Préalables	• Programmation en python • Probabilités et statistiques élémentaires
Thèmes abordés	Le Web Mining est l'application de techniques et modèles permettant de rechercher, récolter, nettoyer, analyser, classer et recommander des informations/données issues du Web. Ces techniques sont notamment utilisées dans les moteurs de recherche, qui jouent un rôle central dans la société de l'information connectée, ainsi que les réseaux sociaux. Ce cours a pour objectif d'apprendre à maîtriser ces techniques et modèles, afin notamment de pouvoir les utiliser/appliquer dans des situations réelles. Les thèmes principaux de ce cours sont : • Recherche d'information sur le Web • Concepts de base • Collecter, nettoyer et analyser les données issues du Web • Text Mining: Analyse de données textuelles issues du Web • Éléments fondamentaux de l'analyse de textes (corpus, bag of words, etc.) • Extraction de termes et représentation de documents (word embedding) • Catégorisation de documents • Analyse • Link Analysis : Analyse de contenu relevant sur base de la structure en hyperliens, formant un réseau/graphe, des données issues du Web • Éléments fondamentaux de la structure d'un réseau/graphe • Méthodologie d'analyse d'un réseau/graphe • Identification de sous-groupes cohésifs • Notions de similarité et de distance • Identification de nœuds prestigieux • Identification de nœuds centraux • Prédiction de nouveaux liens • etc.
Acquis d'apprentissage	A la fin de cette unité d'enseignement, l'étudiant est capable de : <u>Les Acquis d'Apprentissage (AA) au terme de l'unité d'enseignement</u> A la fin du cours l'étudiant sera capable de : 1 • Comprendre le fonctionnement des principaux outils d'extraction de données issues du Web ; • Comprendre le fonctionnement des principaux algorithmes permettant de classer, d'analyser et d'exploiter les informations (sous forme textuelle ou sous forme de liens) issues du Web et utiliser ces algorithmes ; • Prendre les bonnes décisions dans des processus de recherche d'information sur le Web et/ou d'analyse de ces informations.

Modes d'évaluation des acquis des étudiants	• Réalisation d'un projet de groupe; • Défense orale du projet et examen oral, en session. Notes importantes : • Le cours est un cours à évaluation unique (i.e., une fois la note obtenue pour le cours, elle est définitive pour toute l'année académique, et ne peut donc plus être améliorée par la suite) !!! • En soumettant un travail pour évaluation, vous affirmez : (i) qu'il reflète fidèlement le phénomène étudié, et pour cela vous devez avoir vérifié les faits, surtout s'ils sont prétendus par une IA générative (dont vous devez mentionner explicitement l'utilisation en tant qu'outil de soutien à la réalisation de votre travail) ; (ii) avoir respecté toutes les exigences spécifiques du travail qui vous est confié, notamment les exigences pour la transparence et la documentation de la démarche scientifique mise en œuvre. Si l'une de ces affirmations n'est pas vraie, que ce soit intentionnellement ou par négligence, vous êtes en défaut de votre engagement déontologique vis-à-vis de la connaissance produite dans le cadre de votre travail, et éventuellement d'autres aspects de l'intégrité académique, ce qui constitue une faute académique et sera considéré comme telle.
Méthodes d'enseignement	• Exposés magistraux • Exercices intégrés au cours • Projet par groupe
Contenu	Le Web Mining est l'application de techniques et modèles permettant de rechercher, récolter, nettoyer, analyser, classer et recommander des informations/données issues du Web. Ces techniques sont notamment utilisées dans les moteurs de recherche, qui jouent un rôle central dans la société de l'information connectée, ainsi que les réseaux sociaux. Ce cours a pour objectif d'apprendre à maîtriser ces techniques et modèles, afin notamment de pouvoir les utiliser/appliquer dans des situations réelles. Les thèmes principaux de ce cours sont: • Recherche d'information sur le Web • Concepts de base • Modèles de recherche d'information • Collecter, nettoyer et analyser les données issues du Web • Text Mining: Analyse de données textuelles issues du Web • Éléments fondamentaux de l'analyse de textes (corpus, bag of words, etc.) • Extraction de termes et représentation de documents (word embedding) • Catégorisation de documents • Analyse • Link Analysis : Analyse de contenu relevant sur base de la structure en hyperliens, formant un réseau/graphe, des données issues du Web • Éléments fondamentaux de la structure d'un réseau/graphe • Méthodologie d'analyse d'un réseau/graphe • Identification de sous-groupes cohésifs • Notions de similarité et de distance • Identification de nœuds prestigieux • Identification de nœuds centraux • Prédiction de nouveaux liens • etc.
Ressources en ligne	Voir Moodle / Teams
Bibliographie	• MCILWRAITH D., MARMANIS H., BABENKO D. 2nd ed, Algorithms of the Intelligent Web, Manning Publications, 2016. • LANGVILLE A., MEYER C., Google's PageRank and Beyond : The Science of Search Engine Rankings, Princeton University Press, 2012. • FOUSS F., SAERENS M., SHIMBO M., Algorithms and Models for Network Data and Link Analysis, Cambridge University Press, 2016. • AMINI M.-R., GAUSSIER E., Recherche d'information : Applications, modèles et algorithmes, Eyrolles, 2013. • MANNING C. D., RAGHAVAN P., SCHÜTZE H., Introduction to Information Retrieval, Cambridge University Press, 2008. • MARTIN A., CHARTIER M., ANDRIEU O., Techniques de référencement web : Audit et suivi SEO, Eyrolles, 2016.
Autres infos	Certains cours magistraux pourraient être proposées en format hybride co-modal (projection sur le site louvaniste) afin d'élargir et de faciliter la participation des étudiants des deux campus. Les professeurs fourniront des informations supplémentaires à ce sujet en temps voulu.
Faculté ou entité en charge:	CLSM

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Master [120] : ingénieur de gestion	INGE2M	5		
Master [120] : ingénieur de gestion	INGM2M	5		
Master [120] en sciences de gestion (en alternance)	GESA2M	5		