






5.00 crédits	30.0 h + 15.0 h	Q2
--------------	-----------------	----

Enseignants	Nijssen Siegfried ;
Langue d'enseignement	Anglais > Facilités pour suivre le cours en français
Lieu du cours	Louvain-la-Neuve
Thèmes abordés	<p>Une tâche importante dans l'exploration de données consiste à découvrir des motifs dans les données. Ces motifs sont des structures récurrentes dans les données; ils peuvent fournir des explications interprétables pour les observations dans les données, peuvent aider à mieux comprendre la structure des données, peuvent être utilisés pour construire de meilleurs modèles, et peuvent être utilisés pour résoudre d'autres tâches (telles que la constructions d'index dans des bases de données ou la compression de données). Les motifs peuvent être trouvés dans beaucoup de formes de données différentes, y compris des données des supermarchés, des compagnies d'assurance, des expériences scientifiques, des réseaux sociaux, des projets de logiciel, et ainsi de suite.</p> <p>Ce cours offrira une introduction en profondeur à l'extraction de motifs. Après une introduction aux principes de base de l'extraction de motifs, il fournira une discussion en profondeur d'un certain nombre de techniques avancées d'extraction de motif.</p> <p>Les sujets qui seront discutés sont:</p> <ul style="list-style-type: none"> • Catégories de tâches d'extraction de motifs, y compris l'extraction de motifs et d'ensemble de motifs, l'extraction de motifs supervisée ou non, les types de jeux de données et les fonctions de score des motifs ; • Algorithmes pour résoudre différentes tâches d'extraction de motifs; • Les structures de données pour rendre l'extraction des motifs plus efficace; • La mise en 'uvre d'algorithmes d'extraction de motifs ; • Fondements mathématiques pour les différentes catégories de tâches d'extraction de motifs; • Classes de complexité en lien avec l'extraction de motifs ; • Applications de l'extraction de motifs, avec un accent particulier sur l'application des techniques d'extraction de motifs dans l'ingénierie logicielle.
Acquis d'apprentissage	<p>A la fin de cette unité d'enseignement, l'étudiant est capable de :</p> <p>Eu égard au référentiel AA du programme « Master ingénieur civil en informatique », ce cours contribue au développement, à l'acquisition et à l'évaluation des acquis d'apprentissage suivants :</p> <ul style="list-style-type: none"> • INFO 1 • INFO 2.1-4 • INFO 4.2-4 • INFO 5.5 • INFO 6.4 <p>Eu égard au référentiel AA du programme « Master [120] en sciences informatiques », ce cours contribue au développement, à l'acquisition et à l'évaluation des acquis d'apprentissage suivants :</p> <p>1</p> <ul style="list-style-type: none"> • SINF 1.M4, 1.M3 • SINF 2.1-4 • SINF 4.2-4 • SINF 5.5 • SINF 6.4 <p>Les étudiants ayant suivi avec fruit ce cours seront capables de</p> <ul style="list-style-type: none"> • Identifier la tâche d'extraction de motifs la plus appropriée pour un ensemble de données donné; • Expliquer les avantages et les inconvénients des algorithmes d'extraction de motifs en fonction du problème à résoudre; • Identifier les approches appropriées pour évaluer la qualité des motifs et les appliquer dans diverses situations; • Déterminer la complexité calculatoire des problèmes d'extraction de motifs; • Développer de nouveaux algorithmes d'extraction de motifs pour de nouvelles applications.

Modes d'évaluation des acquis des étudiants	<p>The final grade is determined by 3 projects and an exam that is organized at the end of the semester.</p> <p>The grade is calculated following a 75% / 25% rule (final written exam / participation and grade obtained for projects during the semester). Every project counts equally.</p>
Méthodes d'enseignement	<ul style="list-style-type: none"> • Cours magistraux • Travaux pratiques • 3 projets <p>Même si une préférence est donné au séances en présentiel, en fonction de la situation sanitaire et du nombre d'étudiants inscrits une autre forme d'enseignement et de l'évaluation (distanciel, co-modal ou hybride) peut être envisagée.</p>
Contenu	<ul style="list-style-type: none"> • Extraction dans des collections d'objets les plus fréquentes : algorithmes, structures de données; • Extraction dans des collections d'objets basée sur des contraintes: algorithmes, structures de données; • Extraction de motifs dans des séquences, arbres, graphes: algorithmes, structures de données, classes de complexité; • Extraction de motifs dans des données supervisées: fonctions de score, algorithmes; • Extraction d'ensemble de motifs dans des données supervisées: fonctions de score, modèles (arbres de décision, boosting), algorithmes • Extraction d'ensemble de motifs dans des données non supervisées: fonctions de score (principe de longueur de description minimale, entropie maximale), algorithmes • Applications de l'extraction de motifs: dépôts de logiciels, traces, log files, chimio-informatique, bioinformatique, applications industrielles
Ressources en ligne	https://moodle.uclouvain.be/course/view.php?id=3069
Bibliographie	<p>Charu C. Aggarwal, Jiawei Han (Eds.), Frequent Pattern Mining, Springer 2014 (ISBN: 978-3-319-07820-5)</p> <p>Chapitres de Siegfried Nijssen, Albrecht Zimmermann and Luc De Raedt, Essentials of Pattern Mining.</p>
Autres infos	<p>During this course students wil have to implement a number of projects in Python. This course is impossible to follow without prior knowledge of Python. Hence, students should have followed a prior course in Python, such LEPL1401, LINFO1101 or LSINC1101.</p>
Faculté ou entité en charge:	INFO

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Master [120] en science des données, orientation statistique	DATS2M	5		
Master [120] : ingénieur civil en informatique	INFO2M	5		
Master [120] en sciences informatiques	SINF2M	5		
Master [120] : ingénieur civil en mathématiques appliquées	MAP2M	5		
Master [120] : ingénieur civil en science des données	DATE2M	5		
Master [120] en science des données, orientation technologies de l'information	DATI2M	5		