

INF7710

Théorie et applications de la fouille d'associations

Plan de cours

Enseignement

VALTCHEV, Petko
PK-4415
valtchev.petko@uqam.ca
<https://professeurs.uqam.ca/professeur/valtchev.petko/>
Groupes : 040

Description officielle

Description

La découverte d'associations est un aspect fondamental de la fouille de données. Ce cours met l'accent sur les bases théoriques de l'approche et sur les liens avec des problématiques de la théorie de la normalisation en bases de données, l'analyse formelle de concepts et les fonctions Booléennes.

- Problème générique de découverte d'associations et de la fouille de motifs fréquents.
- Variantes : motifs fermés, motifs maximaux, motifs clés ou générateurs.
- Structures algébriques mises en jeux : treillis Booléen, classes d'équivalence, correspondances de Galois, treillis de concepts, contextes.
- Approches de fouille de motifs : algorithmes par niveaux, algorithmes verticaux, algorithmes hybrides.
- Représentations compactes pour les associations : base canonique, bases génériques et informatives.
- Famille réduites de motifs : motifs indérivables, motifs delta-libres, motifs sans disjonction, motifs k-libres.
- Applications de la fouille d'associations.

Objectif du cours

Ce cours vise à approfondir les connaissances de l'étudiant sur un domaine en pleine expansion qu'est la fouille de données. Le cours se focalise sur la découverte d'associations et de motifs fréquents qui est une discipline fondamentale de la fouille de données. L'accent est mis sur la présentation des diverses instanciations du problème général de la fouille ainsi que sur les fondements théoriques de l'approche et leurs liens avec des problématiques de la théorie de la normalisation en bases de données et l'analyse formelle de concepts.

Le but du cours est de permettre à l'étudiant de se familiariser avec la fouille de données en général à travers l'étude plus approfondie d'une des disciplines pertinentes, soit la fouille d'associations. Les objectifs concrets du cours peuvent être énoncés comme suit :

- d'approfondir la maîtrise de certains concepts fondamentaux en bases de données ;
- de familiariser l'étudiant avec la démarche générale en fouille d'associations ;
- de lui fournir des connaissances exploitables en conception de méthodes de fouille ;
- de faire connaître à l'étudiant les plus récents développements dans le domaine ;
- de permettre l'approfondissement d'un des thèmes de recherche dans le domaine ;
- d'initier l'étudiant à la recherche à travers la rédaction d'un rapport sur un sujet d'actualité.

Contenu du cours

Ce cours est destiné aux étudiants ayant suivi au préalable un ou des cours de bases de données au niveau baccalauréat et, de préférence, un cours d'introduction à l'intelligence artificielle. Les sujets abordés sont :

- Introduction à la problématique de la fouille de données
- Principales tâches de la fouille et solutions de principe
- Problème générique de la fouille d'associations et ses défis techniques
- Approche de référence pour l'extraction d'associations
- Approches alternatives et familles réduites de motifs et d'associations
- Ordres partiels et treillis et leur rôle en fouille de motifs
- Analyse de concepts comme cadre fondamental pour fouille de motifs
- Associations, implications, dépendances fonctionnelles
- Motifs et associations structurés : séquences, arbres, graphes
- Fouille en présence de connaissances du domaine : les motifs généralisés
- Principales applications des motifs et associations : médecine, recherche d'information, recommandation, modélisation de l'utilisateur.

Formulé pédagogique et autres modalités d'enseignement

- Séances de **cours magistraux** qui se déroulent, selon la situation sanitaire : a) en ligne (sur Zoom) et en synchrone ou b) en présentiel à l'UQAM.
- Le **lien Zoom** se trouve sur **le site Moodle** du cours
- Les **présentations électroniques** du professeur seront disponibles sur le site Moodle du cours (voir plus bas) *avant* les séances
- **Questions** sur le cours : par courriel ou sur les **forums du site Moodle** (un par semaines)

Modalités d'évaluation

| Description sommaire | Date | Pondération |
|--|------------|-------------|
| Résumé d'article de recherche 1 | Semaine 4 | 15% |
| Résumé d'article de recherche 2 | Semaine 8 | 15% |
| Travail de session : 1re partie - Proposition de sujet | Semaine 11 | 10% |
| Travail de session : 2e partie - Présentation orale | Semaine 15 | 25% |
| Travail de session : 3e partie - Mémoire sur le sujet choisi | Semaine 16 | 35% |

Le travail de session est réalisé individuellement ou par groupe de deux étudiants. Une liste de sujets potentiels et des recommandations seront fournis pendant la session. La qualité du français constitue un critère d'évaluation (pour un maximum de 10%). En cas de retard dans la remise des travaux, une pénalité de 5% par jour ouvrable sera appliquée. Un retard de plus d'une semaine ne sera pas accepté.

Les règlements concernant le plagiat seront strictement appliqués. Pour plus de renseignements, consultez le site suivant : <http://www.sciences.uqam.ca/etudiants/integrite-academique.html>

Matériel requis

En cas d'enseignement à distance : - Ordinateur avec une bonne connexion internet - Casque avec micro - Webcam

Médiagraphie

VC TAN, P.-N., STEINBACH, M., KARPATNE, A. et KUMAR, V. – *Introduction to Data Mining, 2e* – PEARSON (2018)
<http://www-users.cs.umn.edu/~kumar/dmbook/>

VC GODIN, R. – *Systèmes de gestion de bases de données par l'exemple* – LOZE-DION (2006) <http://www.info2.uqam.ca/~godin/livreEd2.html>

VC HAN J., KAMBER, M. et PEI, J. – *Data Mining : Concepts and Techniques, 3e* – MORGAN KAUFMANN (2011).

VC BERRY, M. et LINOFF, G. – *Mastering Data Mining* – JOHN WILEY & SONS (2000).

VC HAND, D., MANNILA, H. et SMYTH, P. – *Principles of Data Mining* – MIT PRESS (2000).

VC CARPINETO, C. et ROMANO, G. – *Concept Data Analysis : Theory and Applications* – WILEY (2004).

VC GANTER, B. et WILLE R. – *Formal Concept Analysis, Mathematical Foundations.* – SPRINGER-VERLAG (1999).

AC FAYYAD, U., PIATETSKY-SHAPIO, G. et SMYTH, P. – *From Data Mining to Knowledge Discovery in Databases* – AI MAGAZINE 17(3) : 37-54, 1996.

AC AGRAWAL, R., IMIELINSKI, T. et SWAMI, A. – *Mining association rules between sets of items in large databases* – PROCEEDINGS OF THE ACM SIGMOD INTERNATIONAL CONFERENCE ON THE MANAGEMENT OF DATA, WASHINGTON (DC), USA, PAGES 207–216, 1993.

AC AGRAWAL, R., MANNILA, H., SRIKANT, R., TOIVONEN, H. et VERKAMO, A. – *Fast Discovery of Association Rules* – U. FAYYAD, G. PIATETSKY-SHAPIO, AND P. SMYTH, ÉDS, ADVANCES IN KNOWLEDGE DISCOVERY AND DATA MINING, PAGES 307–328. AAAI PRESS, MENLO PARK (CA), USA, 1996.

AC D'autres références (selon les sujets abordés) seront remises durant la session.

Information sur les Services à la vie étudiante

Services. Les services à la vie étudiante accompagnent les étudiantes et les étudiants dans la réussite de leur parcours universitaire.

[Ensemble des services offerts](#)

[Gagner du temps et réaliser de meilleurs travaux](#)

[Politiques associées à votre réussite](#)

Bureau. Bureau des services-conseils (**soutien psychologique, bien-être aux études, information scolaire et insertion professionnelle, orientation, emploi**) : pour prendre rendez-vous, communiquez au 514 987-3185 ou par courriel à services-conseil@uqam.ca

Aide financière. Bureau de l'**aide financière** : pour prendre rendez-vous, écrivez à aidefinanciere@uqam.ca

Bourses d'études. Concernant les **bourses**, pensez à consulter Le Répertoire institutionnel des bourses d'études (RIBÉ) et écrivez à bourse@uqam.ca pour toute question.

Informations générales. Consultez les informations et l'ensemble des coordonnées et services offerts par les Services à la vie étudiante à l'adresse suivante : vie-etudiante@uqam.ca.

Politique d'absence aux examens

Reprise d'examen. L'autorisation de reprendre un examen en cas d'absence est de **caractère exceptionnel**. Pour obtenir un tel privilège, l'étudiant.e doit avoir des motifs sérieux et bien justifiés.

Conflits d'horaire. Il est de la responsabilité de l'étudiant.e de ne pas s'inscrire à des cours qui sont en conflit d'horaire, tant en ce qui concerne les séances de cours ou d'exercices que les examens. **De tels conflits d'horaire ne constituent pas un motif justifiant une demande d'examen de reprise.**

Procédure. L'étudiant.e absent.e lors d'un examen doit, dans les cinq (5) jours ouvrables suivant la date de l'examen, présenter une demande de reprise en utilisant le formulaire prévu, disponible sur <https://info.uqam.ca/repriseexamen/>.

Pièces justificatives. Dans le cas d'une absence de moins de cinq (5) jours pour raison médicale, l'étudiant.e doit joindre une déclaration sur l'honneur. Lors d'une absence de cinq (5) jours et plus, un billet médical est exigé. Les dates d'invalidité doivent être clairement indiquées sur le billet original. L'authenticité du billet pourrait être vérifiée.

Dans le cas d'une absence pour une raison non médicale, l'étudiant.e doit fournir les documents originaux expliquant et justifiant l'absence à l'examen ; par exemple, lettre de la Cour en cas de participation à un jury, copie du certificat de décès en cas de décès d'un proche, etc. Toute demande incomplète sera refusée. Si la direction du programme d'études de l'étudiant.e constate qu'un.e étudiant.e a un comportement récurrent d'absence aux examens, l'étudiant.e peut se voir refuser une reprise d'examen.

Pour plus d'informations. Consulter la page <https://info.uqam.ca/repriseexamen/>.

Règlement no 18 sur les infractions de nature académique. (extraits)

[Lien vers la page originale de ce contenu](#)

Tout acte de plagiat, fraude, copiage, tricherie ou falsification de document commis par une étudiante, un étudiant, de même que toute participation à ces actes ou tentative de les commettre, à l'occasion d'un examen ou d'un travail faisant l'objet d'une évaluation ou dans toute autre circonstance, constituent une infraction au sens de ce règlement.

(R18, art. 2.1, définition d'une infraction)

Liste non limitative des infractions mentionnées dans le R18 :

- la substitution de personnes ou l'usurpation d'identité (art. 2.2 a) ;
- le plagiat : l'utilisation totale ou partielle du texte d'autrui ou de la production d'autrui en le faisant passer pour sien ou sans indication de référence (art. 2.2 b) ;
- le recyclage/la réutilisation de travaux : le dépôt d'un travail aux fins d'évaluation alors que ce travail constitue en tout ou en partie un travail qui a déjà été soumis par la personne étudiante, aux fins d'évaluation académique à l'UQAM ou dans une autre institution d'enseignement, sauf avec l'accord préalable de la personne enseignante à qui ce travail est soumis (art. 2.2 c) ;
- la possession ou l'obtention par vol, manœuvre ou corruption de questions ou de réponses d'examen (art. 2.2 d) ;
- la possession ou l'utilisation de tout document ou matériel non autorisé préalablement, pendant un examen ou lors de la réalisation de travaux, incluant le recours aux outils informatiques ou moyens technologiques (art. 2.2 e) ;
- l'utilisation pendant un examen de la copie d'examen ou de tout autre matériel provenant d'une autre personne (art. 2.2 f) ;
- l'obtention de toute aide non autorisée, qu'elle soit collective ou individuelle (art. 2.2 g), ce qui inclut l'utilisation d'outils d'intelligence artificielle ;
- l'obtention d'une évaluation non méritée notamment par corruption, chantage, intimidation ou toute forme de harcèlement ou la tentative d'obtenir une telle évaluation (art. 2.2 h) ;
- la falsification d'un document ou la création d'un faux document, notamment d'un document transmis à l'Université ou d'un document de l'Université transmis ou non à une tierce personne, quelles que soient les circonstances (art. 2.2 i) ;
- la falsification de données de recherche dans un travail, notamment une thèse, un mémoire, un mémoire- création, un rapport de stage ou un rapport de recherche (art. 2.2 j).

Les sanctions reliées à ces infractions sont précisées aux articles 3 et 5 du [Règlement no 18 sur les infractions de nature académique](#).

Pour éviter de vous exposer à des sanctions :

1. Consultez le site r18.uqam.ca pour plus d'information sur l'intégrité académique et le R18 ;
2. Développez les bonnes pratiques en matière de recherche documentaire et de rédaction des travaux via l'outil [Infosphère](#) et les [formations offertes par le Service des bibliothèques](#)

Politique no 2

Le droit à la liberté académique universitaire est le droit de toute personne d'exercer librement et sans contrainte doctrinale, idéologique ou morale, telle la censure institutionnelle, une activité par laquelle elle contribue à l'accomplissement de la mission de l'Université.

Ce droit comprend la liberté :

- (a) d'enseignement et de discussion ;
- (b) de recherche, de création et de publication ;
- (c) d'exprimer son opinion sur la société et sur une institution, y compris l'établissement duquel la personne relève, ainsi que sur toute doctrine, tout dogme ou toute opinion ;
- (d) de participer librement aux activités d'organisations professionnelles ou d'organisations académiques.

Il doit s'exercer en conformité avec les normes d'éthique et de rigueur scientifique généralement reconnues par le milieu universitaire et en tenant compte des droits des autres membres de la communauté universitaire.

En reconnaissant, en promouvant et en protégeant la liberté académique universitaire, cette politique soutient la mission de l'Université, laquelle comprend la production et la transmission de connaissances par des activités de recherche, de création et d'enseignement et par des services à la collectivité.

Pour plus d'information, vous pouvez consulter la section [Liberté académique universitaire](#).

Politique no 16 visant à prévenir et combattre le sexisme et les violences à caractère sexuel

Les violences à caractère sexuel se définissent comme étant des comportements, propos et attitudes à caractère sexuel non consentis ou non désirés, avec ou sans contact physique, incluant ceux exercés ou exprimés par un moyen technologique, tels les médias sociaux ou autres médias numériques. Les violences à caractère sexuel peuvent se manifester par un geste unique ou s'inscrire dans un continuum de manifestations et peuvent comprendre la manipulation, l'intimidation, le chantage, la menace implicite ou explicite, la contrainte ou l'usage de force.

Les violences à caractère sexuel incluent, notamment :

- la production ou la diffusion d'images ou de vidéos sexuelles explicites et dégradantes, sans motif pédagogique, de recherche, de création ou d'autres fins publiques légitimes ;
- les avances verbales ou propositions insistantes à caractère sexuel non désirées ;
- la manifestation abusive et non désirée d'intérêt amoureux ou sexuel ;
- les commentaires, les allusions, les plaisanteries, les interpellations ou les insultes à caractère sexuel, devant ou en l'absence de la personne visée ;
- les actes de voyeurisme ou d'exhibitionnisme ;
- le (cyber) harcèlement sexuel ;
- la production, la possession ou la diffusion d'images ou de vidéos sexuelles d'une personne sans son consentement ;
- les avances non verbales, telles que les avances physiques, les attouchements, les frôlements, les pincements, les baisers non désirés ;
- l'agression sexuelle ou la menace d'agression sexuelle ;
- l'imposition d'une intimité sexuelle non voulue ;
- les promesses de récompense ou les menaces de représailles, implicites ou explicites, liées à la satisfaction ou à la non-satisfaction d'une demande à caractère sexuel.

Pour consulter la politique no 16

https://instances.uqam.ca/wp-content/uploads/sites/47/2019/04/Politique_no_16_2.pdf

Pour obtenir de l'aide, faire une divulgation ou une plainte

Bureau d'intervention et de prévention en matière de harcèlement
514-987-3000, poste 0886

Pour obtenir la liste des services offerts à l'UQAM et à l'extérieur de l'UQAM

<https://harcelement.uqam.ca>

Soutien psychologique (Services à la vie étudiante)

514-987-3185
Local DS-2110

CALACS Trêve pour Elles – point de services UQAM

514 987-0348
calacs@uqam.ca
<http://trevepourelles.org>

Service de la prévention et de la sécurité

514-987-3131

Politique no 44 d'accueil et de soutien des étudiant.e.s en situation de handicap

Politique. Par sa politique, l'Université reconnaît, en toute égalité des chances, sans discrimination ni privilège, aux étudiant.e.s en situation de handicap, le droit de bénéficier de l'ensemble des ressources du campus et de la communauté universitaire, afin d'assurer la réussite de leurs projets d'études, et ce, dans les meilleures conditions possibles. L'exercice de ce droit est, par ailleurs, tributaire du cadre réglementaire régissant l'ensemble des activités de l'Université.

Responsabilité de l'étudiant.e. Il incombe aux étudiant.e.s en situation de handicap de rencontrer les intervenant.e.s (conseiller.ère.s à l'accueil et à l'intégration du Service d'accueil et de soutien des étudiant.e.s en situation de handicap, professeur.e.s, chargé.e.s de cours, direction de programmes, associations étudiantes concernées, etc.) qui pourront faciliter leur intégration à la communauté universitaire ou les assister et les soutenir dans la résolution de problèmes particuliers en lien avec les limitations entraînées par leur déficience.

Service d'accueil et de soutien aux étudiant.e.s en situation de handicap. Le Service d'accueil et de soutien aux étudiant.e.s en situation de handicap (SASESH) offre des mesures d'aménagement dont peuvent bénéficier certains étudiant.e.s. Il est fortement recommandé aux de se prévaloir de ces services afin de réussir ses études, sans discrimination. Pour plus d'information, visiter le site de ce service : <https://services.uqam.ca/services-offerts/soutien-aux-etudiants-en-de-situation-handicap/> et celui de la politique institutionnelle d'accueil et de soutien aux étudiant.e.s en situation de handicap : https://instances.uqam.ca/wp-content/uploads/sites/47/2018/05/Politique_no_44.pdf

Il est important d'informer le SASESH de votre situation le plus tôt possible :

- En personne : 1290, rue Saint-Denis, Pavillon Saint-Denis, local AB-2300
- Par téléphone : 514 987-3148
- Par courriel : situation.handicap@uqam.ca
- En ligne : <https://vie-etudiante.uqam.ca/>

Politique no 42 sur le harcèlement

L'Université du Québec à Montréal (ci-après, l'« Université ») reconnaît à toutes les personnes membres de la communauté universitaire le droit d'être traitées avec dignité, équité et respect mutuel.

Toutes, tous sont susceptibles de subir du harcèlement. L'Université reconnaît que le harcèlement est majoritairement dirigé à l'endroit de certains groupes. Il s'agit notamment des femmes, plus particulièrement lorsque leur vécu se situe à l'entrecroisement de plusieurs formes de discrimination, des personnes issues des minorités sexuelles ou de genre, des communautés racisées ou ethniciées, des communautés autochtones, des étudiantes, étudiants internationaux, ainsi que des personnes en situation de handicap. L'Université s'engage donc à tenir compte de leurs besoins spécifiques.

L'Université considère le respect mutuel, l'égalité, l'écoute et l'entraide comme des valeurs importantes qui favorisent l'épanouissement personnel ainsi que l'établissement de rapports harmonieux entre les personnes et entre les groupes, et qui permettent la mise en place d'un milieu sain et propice à la réalisation individuelle ou collective de sa mission universitaire.

L'Université croit que la collaboration de chaque personne et de chaque groupe de la communauté universitaire est essentielle pour favoriser la création d'un tel milieu et, en ce sens, elle compte sur la contribution de chaque personne.

L'Université juge que toute forme de harcèlement porte atteinte à la dignité et à l'intégrité physique ou psychologique d'une personne.

L'Université reconnaît sa responsabilité d'assurer un milieu de travail et d'études exempt de toute forme de harcèlement et veille à ce qu'aucune forme de harcèlement ne soit tolérée, quelle qu'en soit la source.

Pour plus de détails, consultez la politique complète : https://instances.uqam.ca/wp-content/uploads/sites/47/2018/05/Politique_no_42.pdf

Monitorat de programme

Le département d'informatique offre un service gratuit d'aide à la réussite s'adressant plus particulièrement aux étudiant.e.s du baccalauréat et du certificat en informatique. Il concerne principalement les cours de base comme INF1070, INF1120, INF1132, INF2120 et INF2171, mais, selon la connaissance du moniteur ou de la monitrice, un support dans d'autres cours peut également être offert.

Objectifs. Permettre aux étudiant.e.s de :

- Bénéficier d'un encadrement par les pairs ;
- Recevoir un suivi personnalisé en cas de difficulté ;
- Profiter d'un soutien supplémentaire à la matière vue en classe ;
- Obtenir un support technique sur les technologies, les outils, les bibliothèques et les cadriciels utilisés dans les cours (installation, configuration, utilisation)

Informations.

- Voir <https://info.uqam.ca/aide/> pour la grille horaire et tous les détails
- Le service est généralement disponible à partir de la deuxième semaine
- D'autres plages horaires pourraient être ajoutées en cours de session selon les besoins
- Clavardage en direct : [~aide](#) (Mattermost)