

Reconnaissance des formes

Groupe 30

Mercredi, de 13h30 à 16h30 PK-4360 (cours)

Responsable(s) du cours

Nom du coordonnateur : BOUGUessa, Mohamed**Nom de l'enseignant :** BOUGUessa, Mohamed**Local :** PK-4915**Téléphone :** (514) 987-3000 #5541**Courriel :** bouguessa.mohamed@uqam.ca**Site Web :** <https://sites.google.com/site/mohamedbouguessa/>

Description du cours

Base mathématique pour la reconnaissance des formes, méthodes syntaxiques et statistiques, applications en reconnaissance de la parole et reconnaissance de caractères.

Préalables académiques :

DIC9250 Génie cognitif

Objectifs du cours

Que vise ce cours? Ce cours vise à approfondir les connaissances de base de l'étudiant dans les domaines de forage de données (data mining) ainsi que l'apprentissage automatique (machine learning) en le familiarisant avec les travaux de recherche dans ces domaines.

Les objectifs spécifiques de ce cours :

- Comprendre le processus d'extraction des connaissances dans les bases de données;
- Se familiariser avec les techniques de forage de données prédictives et descriptives;
- Introduire des concepts avancés en forage de données : données non-structurées, gestion de masse de données.
- Travailler sur un projet de recherche qui se rattache à l'intelligence artificielle, l'apprentissage automatique, forage de données, big data, etc.

Contenu du cours

Concepts de base et définitions

- Classification et régression;
- Frontière de décision et fonctions discriminantes;
- Fonction d'erreur et risque empirique;
- Concepts d'apprentissage non supervisé;
- Malédiction de la dimension.

Arbres de décision

- Algorithmes : ID3, C4.5;
- Techniques pour traiter le sur-apprentissage.

Approches bayésiennes

1. Inférence bayésienne;
2. Les K-plus proches voisins.

Machines à noyaux

- SVM (Support Vector Machines).

Apprentissage non supervisé

- Clustering par partition et hiérarchique;
- Clustering basé sur la densité;
- Clustering de graphes;
- Les mélanges de loi de distribution statistiques;
- Techniques de réduction de la dimension;
- Clustering dans les sous-espaces de dimensions : Projected Clustering.

Texte mining

- Approches pour l'analyse et la classification automatique des documents.

Web mining

- Principes et éléments de bases;
- Algorithmes de classement des pages Web : HITS et PageRank de Google.

Formules pédagogiques

- Le cours sera donné sous forme d'exposés magistraux.
- Des périodes de consultation seront déterminées en classe.

Modalités d'évaluation

- Examen final 46%
 - Travail pratique 12%
 - Projet de session 42%
-

Politique d'absence aux examens

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.

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.**

Dans le cas d'une absence pour raison médicale, l'étudiant-e doit joindre un certificat médical original et signé par le médecin décrivant la raison de l'absence à l'examen. Les dates d'invalidité doivent être clairement indiquées sur le certificat. Une vérification de la validité du certificat pourrait être faite. 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 étudiant a un comportement récurrent d'absence aux examens, l'étudiant-e peut se voir refuser une reprise d'examen.

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 le site Web du département à l'adresse suivante : <http://info.uqam.ca/politiques/>

L'étudiant-e doit déposer le formulaire dûment complété au secrétariat de la direction de son programme d'études : PK-3150 pour les programmes de premier cycle, PK-4150 pour les programmes de cycles supérieurs. Pour plus de détails sur la politique d'absence aux examens du Département d'informatique, consultez le site web suivant : <http://info.uqam.ca/politiques>

Intégrité académique

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

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.

La liste non limitative des infractions est définie comme suit :

- la substitution de personnes;
- l'utilisation totale ou partielle du texte d'autrui en la faisant passer pour sien ou sans indication de référence;
- la transmission d'un travail pour fins d'évaluation alors qu'il constitue essentiellement un travail qui a déjà été transmis pour fins d'évaluation académique à l'Université ou dans une autre institution d'enseignement, sauf avec l'accord préalable de l'enseignante, l'enseignant;
- l'obtention par vol, manoeuvre ou corruption de questions ou de réponses d'examen ou de tout autre document ou matériel non autorisés, ou encore d'une évaluation non méritée;
- la possession ou l'utilisation, avant ou pendant un examen, de tout document non autorisé;
- l'utilisation pendant un examen de la copie d'examen d'une autre personne;
- l'obtention de toute aide non autorisée, qu'elle soit collective ou individuelle;
- la falsification d'un document, notamment d'un document transmis par l'Université ou d'un document de l'Université transmis ou non à une tierce personne, quelles que soient les circonstances;
- la falsification de données de recherche dans un travail, notamment une thèse, un mémoire, un mémoire-crédation, un rapport de stage ou un rapport de recherche;
- Les sanctions liées à ces infractions sont précisées à l'article 3 du Règlement no 18.

Les règlements concernant le plagiat seront strictement appliqués. Pour plus de renseignements, veuillez consulter les sites suivants : <http://www.sciences.uqam.ca/etudiants/integrite-academique.html> et <http://www.bibliotheques.uqam.ca/recherche/plagiat/index.html>

Médiagraphie

1. R. O. Duda, P. E. Hart, D. G. Stork. Pattern Classification (2nd ed.), Addison Wesley, 2006.
2. Tom M. Mitchell. Machine Learning, McGraw Hill, 1997.
3. P-N Tan, M. Steinbach, V. Kumar. Introduction to Data Mining, Addison Wesley, 2006.
4. A. Cornuéjols, L. Miclet, Y. Kodratov. Apprentissage artificiel: Concepts et algorithmes, Eyrolles, 2002.

5. J. S-Taylor, N. Cristianini. Support Vector Machines, Cambridge University Press, 2000.

A : article - C : comptes rendus - L : logiciel
S: Standard - U : uri - V : volume

C : complémentaire - O : Obligatoire - R : recommandé