
PLAN DE COURS

Forage de données -Data mining-
(8INF954)

Département d'informatique et de mathématique

Formule pédagogique

Les cours magistraux sont dispensés lors de la période réservée au cours et spécifiée à l'horaire de cours officiel.

Insertion du cours dans le programme

Le cours est optionnel dans les divers programmes d'études de cycles supérieurs du département d'informatique et de mathématique (DIM) de l'Université. Ce cours (8INF954) est destiné aux étudiants ayant un bac. en informatique. Il ne nécessite pas de préalables, mais il est souhaitable d'avoir une bonne base en informatique et en mathématique-statistiques.

Objectifs généraux du cours

Révolution annoncée de l'ingénierie de la connaissance, le forage de données -data mining ou fouille de données- promet de devenir l'outil de veille technologique par excellence. À cause de la croissance grandissante des banques de données et du Web, le développement des méthodes automatiques pour la découverte de connaissances, la classification et la prédiction est devenu indispensable dans de nombreuses applications telles que la finance, la sécurité publique, la santé, etc. Ce cours vise à introduire les étudiants aux techniques de data mining. À la fin du cours l'étudiant sera capable de mener des projets de data mining.

Objectifs spécifiques du cours

Le cours s'inscrit dans une problématique générale liée à l'extraction de connaissances et présente les concepts fondamentaux, ainsi que les techniques principales du data mining. Au terme de ce cours, les étudiants devront être en mesure:

- (i) de maîtriser le processus général de data mining;
- (ii) de distinguer les avantages et les inconvénients des principales approches;
- (iii) de se familiariser avec les algorithmes de data mining.

Contenu du cours

Cours magistraux

Le cours permettra aux étudiants l'apprentissage des éléments suivants (voir l'annexe 1 pour plus de détails) :

1. Concept de machine d'apprentissage par les exemples;
2. Algorithmes de classification, de clustering et de schéma d'association ou de séquence de motifs fréquents;
3. Évaluation de la crédibilité d'un modèle d'apprentissage- data mining-;
4. Ingénierie de la connaissance et processus de data mining.

Références

Ian H. Witten & Eibe Franck, data mining: practical machine learning tools and techniques, Elsevier editor, 2014.

Stéphane Tufféry, data mining et statistique décisionnelle, ISBN : 978-2-7108-0888-6, éditions technip, 2007

Daniel T. Larose, Data mining Methods and Models, Wiley pub., 2009.

Quelques sites intéressants :

L'environnement de la machine d'apprentissage Weka

<http://www.cs.waikato.ac.nz/ml/weka/>

Le site du cours

<http://depcom.uqac.ca/~abouzoua/>