

DIC9370 – Reconnaissance de formes

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

- Inférence bayésienne;
- 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.