Discipline(s) : Infomatique et télécommunications

DATA MINING AND VISUALIZATION

Nature	UE

RESPONSABLES

Alexandre Termier

OBJECTIFS

L'objectif de ce cours est de présenter aux étudiants des notions de fouille de données exploratoire. Cette discipline part de données où des connaissances potentiellement utiles sont présentes, mais on ne sait pas exactement quoi chercher au départ. C'est une situation très commune en science et dans l'industries. Les techniques présentées sont le pattern mining (découverte de différentes formes de régularité dans les données), le data mining déclaratif (méthode simplifiée pour définir et raffiner ce que l'on cherche) et le data mining interactif (co-construction de solution entre l'algorithme et l'utilisateur). Par ailleurs, le cours présente des notions de base de visualisation d'information, qui sont indispensables pour présenter les résultats à un analyste et l'aider à interagir avec le système.

MOTS-CLÉS

Pattern mining, frequent itemset mining, frequent subgraph mining, patternset mining, generic and declarative data mining, interactive data mining, data visualization

PRÉREQUIS

Connaissances et intérêt pour l'algorithmique.

CONTENU

- Méthodologie de KDD
- Les différents types de données
- Prétraitements
- Pattern mining
  - Itemsets fréquents
  - Règles d’association
  - Mesures d’intérêt
  - Fermeture de la connexion de Galois
  - Prise en compte de hiérarchies dans les données
  - Séquences fréquentes
  - Sous-graphes fréquents
  - Patterns discriminants
  - Patterns sets : approches par optimisation, approche par dominance (Skypatterns)

Performance de calcul: approches parallèles
Data mining déclaratif
But: faciliter l'intégration de connaissances du domaine
Algorithmes généraux
Approches à base de contraintes
Autres formalismes: BD Inductives, ASP
Data mining interactif
Applications : fouille de texte, données de supermarché, données de jeux vidéos populaires

COMPÉTENCES ACQUISES

Savoir: algorithmes de pattern mining, méthodes de data mining déclaratif et interactif, visualisation.
Savoir-faire: formaliser une hypothèse d'analyse des données, l'exprimer de façon déclarative, choisir le bon algorithme ou outil pour faire le calcul. Choisir la bonne visualisation pour un problème donné.

APPARTIENT À

Master 2 informatique parcours Science Informatique

Mise à jour le 17 juillet 2017

CONTACT(S)

Département Informatique et télécommunications
École normale supérieure de RennesCampus de Ker LannAvenue Robert Schuman
35170 BRUZ
Tél. : 02 99 05 52 43
E-mail
Site Internet