

Master Recherche Informatique

Apprentissage et Fouille de Données

Michèle Sebag, Antoine Cornuéjols, Balazs Kegl

`http://tao.lri.fr/`

Pourquoi ?

- Prédiction
pannes, maladies, achats,...
- Identification de modèles
ex: facteurs de risque, maladies cardio-vasculaires
- Recherche sur le Web
que serait Super-Google ?
- Optimisation—Conception
Comment minimiser le taux de pannes ?

Data Mining – Fouille de Données

Parmi les 10 technologies émergentes du 21^e siècle (MIT Review, 2001)

Fouille de Données – Motivations

L'idéal le siècle des connaissances

La réalité des expertises spécialisées
des pratiques à l'état de traces dans les données

Le besoin la gestion humaine des connaissances
ne passe pas à l'échelle

L'opportunité les données sont accessibles

OBJECTIF

restituer à l'expert
des connaissances nouvelles, utiles, valides

Objectifs et prérequis

Intelligibilité

Règles; attributs pertinents
Esprit de finesse

Qualité de prédiction

Optimisation: apprentissage statistique; algorithmes génétiques
Pas d'aversion pour les maths

Examen

Questions de cours
Résumé et présentation d'un article récent

Exemples de stages

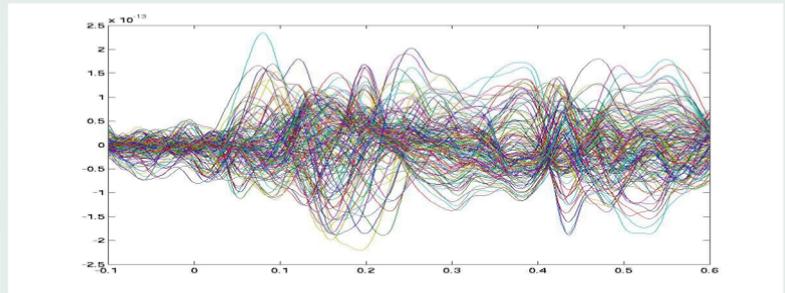
- Fouille de données médicales
- Robotique
- Imagerie cérébrale
- Apprentissage pour les jeux
- Autonomic Computing

Imagerie Cérébrale

Dispositif



Données



Bonnes adresses

- SIGKDD Explorations

[//www.acm.org/sigs/sigkdd/explorations](http://www.acm.org/sigs/sigkdd/explorations)

- CACM Special Issues on Data Mining (Nov 96):

[//www.digimine.com/usama/datamine/acm-contents.htm](http://www.digimine.com/usama/datamine/acm-contents.htm)

- Kdnuggets:

[//www.kdnuggets.com/](http://www.kdnuggets.com/)

- SIGMOD

[//www.acm.org/sigmod/](http://www.acm.org/sigmod/)

- VLDB

[//www.vldb.org](http://www.vldb.org)

- SIGMOD Record:

[//www.acm.org/sigmod/record](http://www.acm.org/sigmod/record)

- IEEE Data Engineering Bulletin:

[//www.research.microsoft.com/research/db/debull](http://www.research.microsoft.com/research/db/debull)