

# Plan cours

## App. et fouille de données (2006-2007)

(10-12-2006)

### Préambule sur le cours :

**8 cours**

#### Documents :

- Transparents et informations sur le site
- Livre « Apprentissage artificiel. Concepts et algorithmes ». A. Cornuéjols & L. Miclet. Eyrolles. 2002.

#### Travaux :

- Contrôle de connaissances sur table
- Exposé sur article

## 1. Introduction et généralités sur l'apprentissage artificiel

### 1.1. Visages de l'apprentissage

#### 1.1.1. Des exemples

- 1.1.1.1. D'applications
- 1.1.1.2. D'apprentissages naturels

#### 1.1.2. Objectifs de l'AA

- 1.1.2.1. Motivations pratiques

**Aspect central : généraliser à partir de cas particuliers (en particulier l'induction)**

- 1.1.2.2. Motivations scientifiques

#### 1.1.3. Des scénarios

#### 1.1.4. Ingrédients de l'apprentissage

- 1.1.4.1. Espace  $H$
- 1.1.4.2. Espace  $X$
- 1.1.4.3. Fonction de perte (ou mesure de performance)
- 1.1.4.4. Méthode d'exploration de  $H$  et d'optimisation
- 1.1.4.5. Protocole

#### 1.1.5. Niveaux d'analyse de l'apprentissage

- 1.1.5.1. Analyse de principe, de faisabilité  
Que peut-on apprendre ? Sous quelles conditions ?
- 1.1.5.2. Niveau de réalisation de systèmes apprenants  
Comment apprendre ? (algorithmes)

### 1.1.6. Types d'apprentissage

- 1.1.6.1. Supervisé
- 1.1.6.2. Non supervisé
- 1.1.6.3. Par renforcement

## 1.2. Des algorithmes

- 1.2.1. PPV
- 1.2.2. Réseaux de neurones : le perceptron
- 1.2.3. Arbres de décision

## 1.3. Approche plus formelle de l'induction

- 1.3.1. L'apprentissage comme un jeu entre espaces
  - 1.3.1.1. Scénario fondamental
  - 1.3.1.2. Importance de l'hypothèse i.i.d.
- 1.3.2. Utilité d'un espace d'hypothèses
- 1.3.3. Choix de l'espace des hypothèses (et langage de description de H)
- 1.3.4. Questions fondamentales
  - 1.3.4.1. Objectif de l'apprentissage : risque réel
  - 1.3.4.2. Choix du critère inductif
  - 1.3.4.3. Choix de H
  - 1.3.4.4. Méthode d'exploration de H
- 1.3.5. Critères inductifs
- 1.3.6. Conditions sur l'induction : approche statistique
- 1.3.7. Une grande typologie des algorithmes d'apprentissage

## 1.4. Apprentissage par exploration

- 1.4.1. Traduction du principe de MRE : cohérence
  - 1.4.1.1. Complétude
  - 1.4.1.2. Correction
- 1.4.2. Notion de relation de généralité dans H
- 1.4.3. Espace des versions et bornes
- 1.4.4. L'algorithme d'élimination des candidats
- 1.4.5. Illustration et TD

- 2. Apprentissage statistique**
  - 2.1. Les séparateurs à Vastes Marges – Méthodes à noyaux
  - 2.2. Analyse de Vapnik
  - 2.3. La sélection de modèles
- 3. Apprentissage par méthodes d'ensembles**
  - 3.1. Bagging
  - 3.2. Boosting
- 4. Apprentissage non supervisé**
  - 4.1. PCA – k-moyennes
  - 4.2. Isomap / Random projections
  - 4.3. One-class SVM
- 5. Réseaux de neurones / Validation de l'apprentissage**
- 6. Apprentissage de séquences et en grande dimension**
  - 6.1. Data streaming
  - 6.2. Apprentissage incrémental et en-ligne
  - 6.3. Apprentissage actif
  - 6.4. Sélection d'attributs
- 7. Fouille de données**
  - 7.1. Le problème
  - 7.2. Text mining
  - 7.3. Web Mining
- 8. Fouille de données (suite et fin)**
  - 8.1. Apprentissage relationnel
  - 8.2. Noyaux relationnels – Applications à la bio-informatique
  - 8.3. Apprentissage collaboratif – Apprentissage de tri