

## SI221 Bases de l'Apprentissage/ Machine learning for pattern recognition

### Introduction

Septembre 2018

Laurence Likforman-Sulem  
Telecom ParisTech/IDS Image Données Signal  
bureau : E 504  
likforman@telecom-paristech.fr

L. Likforman - Telecom ParisTech - SI 221

1

---

---

---

---

---

---

### Equipe pédagogique



- Laurence Likforman, Chloé Clavel (IDS) coordinatrices
- Giovanna Varni, Hamid Jalalzai (doctorant IDS)

L. Likforman - Telecom ParisTech - SI 221

2

---

---

---

---

---

---

### Quelques formalités ....

- Présence obligatoire aux TPs
- Contrôle écrit : 1h30 heures
  - voir site pédagogique (Documents en Ligne)

L. Likforman - Telecom ParisTech - SI 221

3

---

---

---

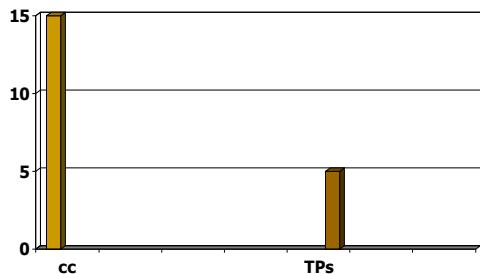
---

---

---

## notation

CC: 15 points , TPs: 5 points



L. Likforman - Telecom ParisTech - SI 221

4

## Site pédagogique

- <https://sitepedago.telecom-paristech.fr>
  - Bases de l'apprentissage
  - documents en ligne, liens, exercices corrigés

L. Likforman - Telecom ParisTech - SI 221

5

## Domaines connexes

- intelligence artificielle, *Apprentissage statistique et fouille de données, Machine learning avancé, Machine Learning for Text Mining, UEs MDI343, INFMDI341, SD-TSIA214*
- probabilités, statistiques, méthodes neuronales, apprentissage profond (Deep Learning)
- sciences cognitives, neurosciences, psychologie perceptive

L. Likforman - Telecom ParisTech - SI 221

6

## Apprentissage/Reconnaissance des formes

- en anglais
  - *machine learning*
  - *pattern recognition*
  - *pattern classification*

L. Likforman - Telecom ParisTech - SI 221

7

---

---

---

---

---

---

---

## Reconnaissance des formes

### ■ Parole

- reconnaissance du locuteur, reconnaissance prise de parole
- reconnaissance de la parole
- reconnaissance des émotions

### ■ Document/Ecriture

- OCR : Optical Character Recognition
- reconnaissance de l'écriture (chèques, enveloppes, lettres, documents historiques)
- identification-authentification de scripteur
- camera OCR : lecture de plaques, de panneaux....
- classification automatique de textes
- authentification d'auteur (authorship verification)
- détection de troubles cognitifs

L. Likforman - Telecom ParisTech - SI 221

8

---

---

---

---

---

---

---

## Reconnaissance des formes

### ■ Vision

- suivi d'objets, de personnes, attitude
- reconnaissance de visages, de sourires
- reconnaissance d'expressions faciales, des émotions (joie, tristesse, surprise, peur, colère et dégoût)
- contrôle de qualité
- interprétation d'images satellites

### ■ Audio

- reconnaissance de notes de musique, d'instruments

### ■ Biométrie

- visages 2D, 3D
- forme/texture/lignes de la main
- empreintes digitales
- multimodalité : voix-visage, main-empreinte,...

L. Likforman - Telecom ParisTech - SI 221

9

---

---

---

---

---

---

---

### robot "taille de vigne"



L. Likforman - Telecom ParisTech - SI 221

10

---

---

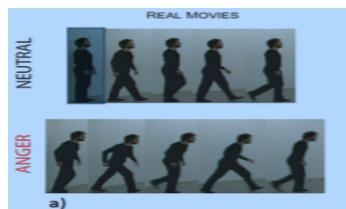
---

---

---

---

### reconnaissance d'émotion à partir de l'attitude



d'après [Hicheur 2013]

L. Likforman - Telecom ParisTech - SI 221

11

---

---

---

---

---

---

### Polémique Molière-Corneille



- Caractéristiques : mots employés
- (mais dépend du sujet)

L. Likforman - Telecom ParisTech - SI 221

12

---

---

---

---

---

---

## Polémique Molière-Corneille

### ■ exemples de caractéristiques

- X02 nombre de phrases simples d'une proposition
- X04 nombre de propositions coordonnées dans une phrase
- X21 nombre de verbes conjugués dans une phrase
- X31 nombre de sujets dans une phrase
- X32 nombre de pronoms sujets dans une phrase

---

---

---

---

---

---

---

## Nouveau poème de Shakespeare ?

[Ian H. Witten, 2005]

Shall I die? Shall I fly  
Lovers' baits and deceits,  
sorrow breeding?  
Shall I tend? Shall I send?  
Shall I sue, and not rue  
my proceeding?  
In all duty her beauty  
binds me her servant for ever,  
If she scorn, I mourn,  
I retire to despair, **joying** never.

*nouveaux mots*  
admiration  
besots  
exiles  
inflection  
**joying**  
scanty  
speck  
tormentor  
twined

■ 430 words in all

---

---

---

---

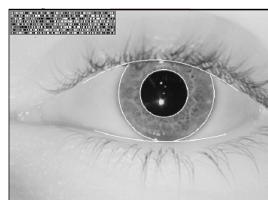
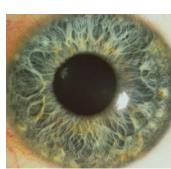
---

---

---

## Biométrie : image de l'iris

[Daugman, 2005]



Texture unique

---

---

---

---

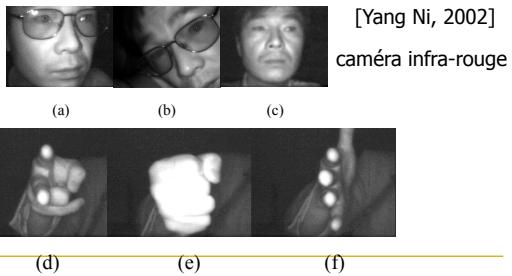
---

---

---

## Projet GET-BIOMET (2001-2002)

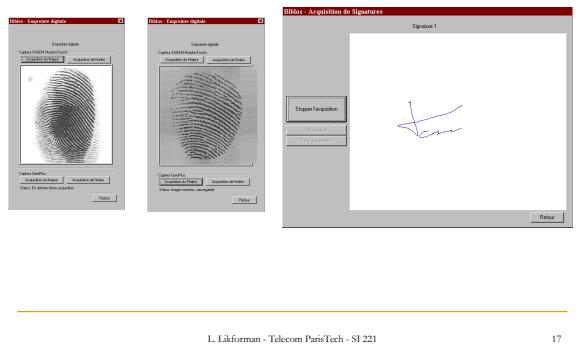
Base : 100 enregistrements x 3 sessions



L. Likforman - Telecom ParisTech - SI 221

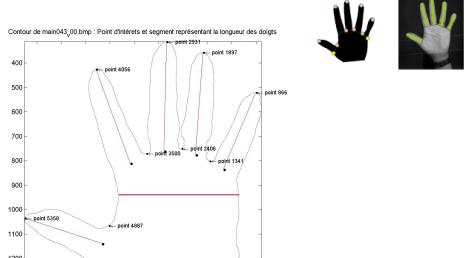
16

## empreintes digitales - signatures



17

## Géométrie de la main

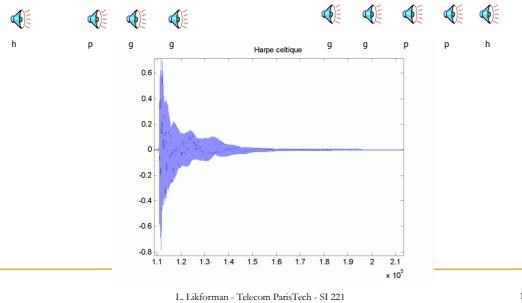


L. Likforman - Telecom ParisTech - SI 221

18

## reconnaissance d'instruments

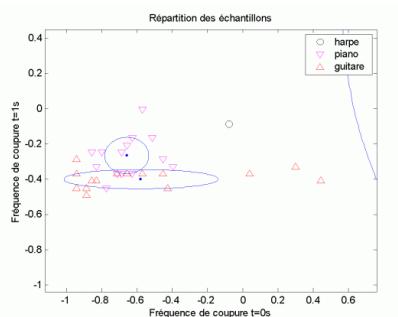
piano, guitare, harpe celtique



L. Likforman - Telecom ParisTech - SI 221

19

Caractéristiques :  
fréquences de coupure (sur enveloppe spectrale)



20

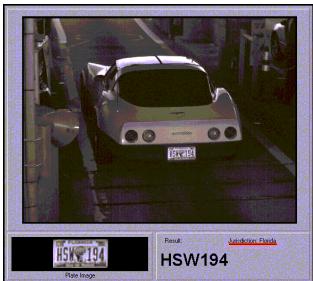
## lecture optique de panneaux



L. Likforman - Telecom ParisTech - SI 221

21

## sécurité routière



L. Likforman - Telecom ParisTech - SI 221

22

---

---

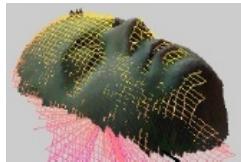
---

---

---

---

## reconnaissance 3D



[Lu et al., 2003]

Xiaoguang Lu, Dirk Colby, and Anil K. Jain, Three-Dimensional Model Based Face Recognition, ICPR 2004, Cambridge

L. Likforman - Telecom ParisTech - SI 221

23

---

---

---

---

---

---

## Objectif de l'apprentissage

Soit  $R$  l'ensemble de représentation, on cherche une fonction de décision  $d$

$$d : R \rightarrow J$$
$$x \mapsto i$$

$J$  = ensemble des étiquettes des classes

$x$  : représentation d'une forme (vecteur, chaîne, séquence,...)

L. Likforman - Telecom ParisTech - SI 221

24

---

---

---

---

---

---

## Difficulté : variabilité

- variabilité inter-classes

S S 7 7 5 5 9



- variabilité intra-classe

4 4

L. Likforman - Telecom ParisTech - SI 221

25

---

---

---

---

---

---



L. Likforman - Telecom ParisTech - SI 221

26

---

---

---

---

---

---

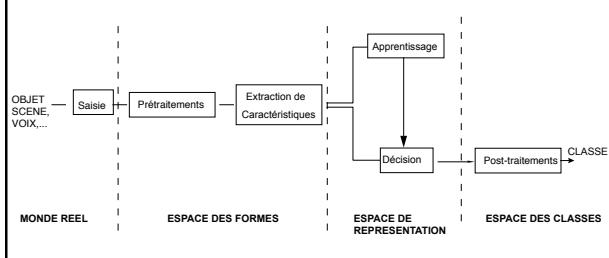
---

---

---

---

## Chaîne des traitements



L. Likforman - Telecom ParisTech - SI 221

27

---

---

---

---

---

---

---

---

---

---

## prétraitements

parer de deux Ponts per  
cord. Près de Lexington, c  
verent un gros d'habitans



input image  
verso apparaissant  
sur le recto

image filtrée

parer de deux Ponts per  
cord. Près de Lexington, c  
verent un gros d'habitans

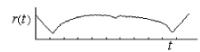
L. Likforman - Telecom ParisTech - SI 221

28

## représentations

### ▪ signature polaire

- $r(t)$ : distance entre points du contour et centre de gravité



[D. Zhang, G. Lu, 2003]

### ▪ moments $m_{pq}$ ou $\mu_{pq}$

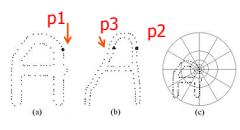
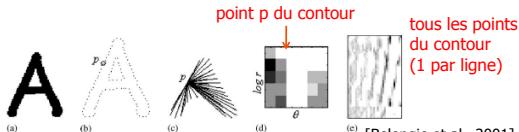
- $i_{mean}, j_{mean}$ : coordonnées centre de gravité,
- $I(i,j)=1$  si pixel noir, 0 sinon

$$m_{pq} = \sum_j \sum_i (i^p j^q * I(i, j))$$

$$\mu_{pq} = \sum_j \sum_i (i - i_{mean})^p (j - j_{mean})^q I(i, j).$$

29

## représentation: shape context



source: Wikipedia

L. Likforman - Telecom ParisTech - SI 221

30

## tf-idf représentation

- V représente document d
    - $V=(a_1, a_2, \dots, a_m)$  avec  $a_i$ : mesure tf-idf terme  $t_i$
  - on veut tenir compte de:
    - occurrence terme  $t_i$  dans document d
    - occurrence terme  $t_i$  dans collection D de N documents
  - meilleurs termes
    - fréquents dans document d
    - peu fréquents dans collection D de documents
    - $a_i = \text{term frequency} * \text{inverse document frequency}$
- $tf(t_i, d) = \text{nombre de fois où } t_i \text{ apparaît dans document } d$  (normalisé par #mots de d)
- $$idf(t_i) = \log \left[ \frac{1 + N}{1 + \# \text{documents including } t_i} \right]$$
- $$a_i = tf(t_i, d) * idf(t_i)$$

L. Likforman, Telecom ParisTech,

31

## exemple

- poids tf pour terme « handwriting » apparaissant 3 fois dans un document d de 100 mots
  - $tf(\text{« handwriting »}, d) = 3/100 = 0.03$
- poids idf pour terme « handwriting » apparaissant dans 1000 documents d'une collection de 10 000 000 documents
  - $idf(\text{« handwriting »}) = \log(10\ 000\ 001/1001)$
- $tf(\text{« handwriting »}, d) * idf(\text{« handwriting »})$   
 $= 0.03 * 9.21 = 0.27$

L. Likforman, Telecom ParisTech,

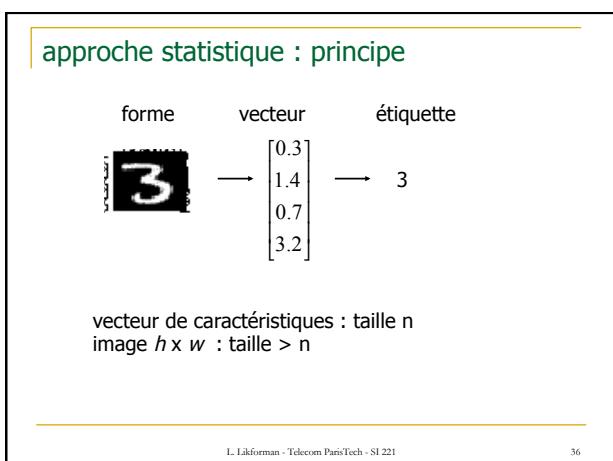
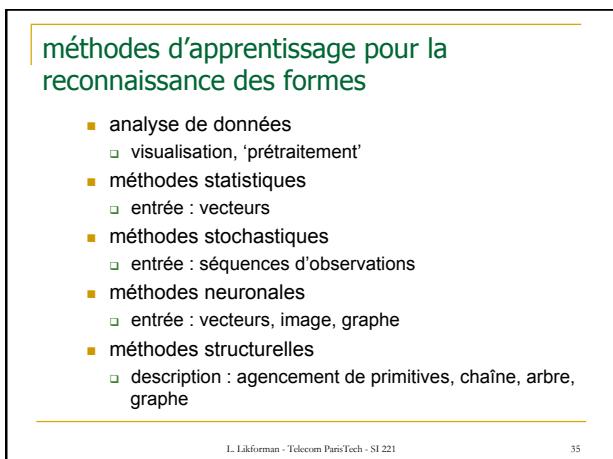
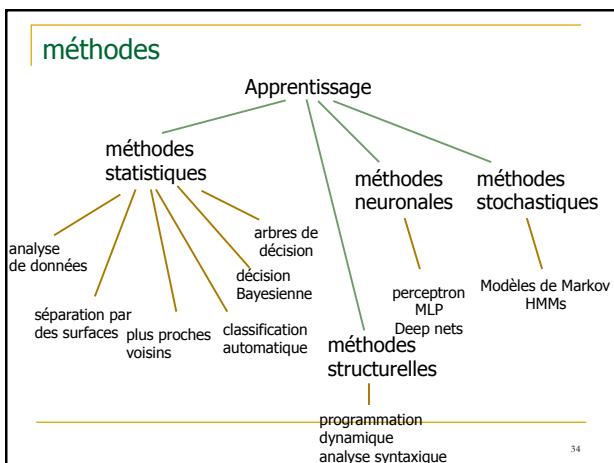
32

## normalisation des données

- représentation vectorielle
  - $x^{(1)}, x^{(2)}, \dots, x^{(N)}$  N vecteurs d'apprentissage
  - $x^{(k)}$ : vecteur de d composantes  $x^{(k)} = x_1^{(k)}, x_2^{(k)}, \dots, x_d^{(k)}$
  - $(k=1, \dots, N)$
- indice RMS: root mean square
  - $$RMS(i) = \sqrt{\frac{\sum_{k=1}^N (x_i^{(k)})^2}{N}}$$
- composante normalisée
  - $$x_i^{(k)} \leftarrow \frac{x_i^{(k)}}{RMS(i)}$$

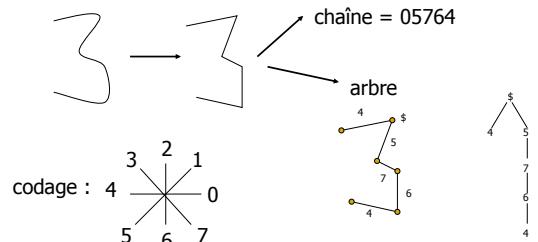
L. Likforman, Telecom ParisTech,

33



## Approche structurelle

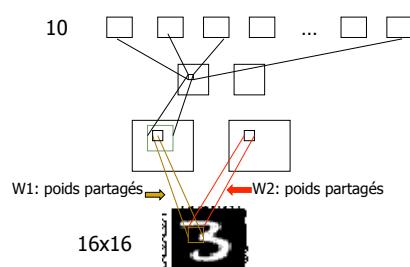
représentations : chaînes, arbres , graphes



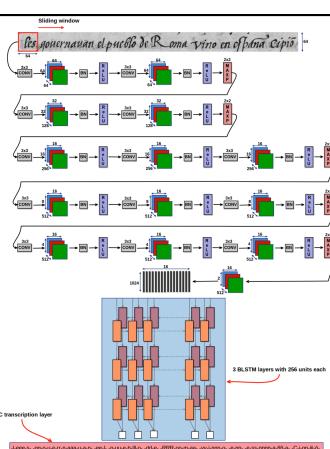
## Approches neuronales

Le-Net [LeCun et al.]

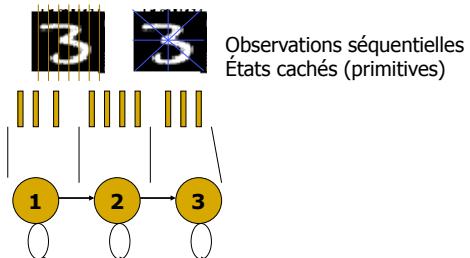
réseaux convolutionnels



## Convolutional RNN



## Approches Markoviennes



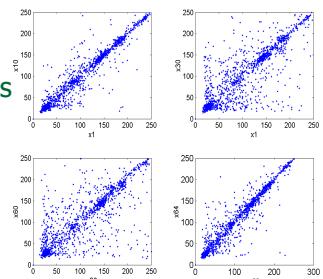
....extension : Réseaux Bayésiens Dynamiques

L. Likforman - Telecom ParisTech - SI 221

40

## analyse de données

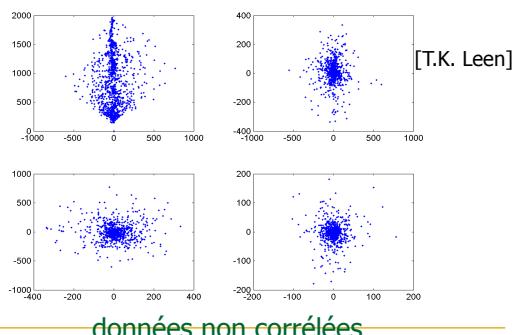
données corrélées



L. Likforman - Telecom ParisTech - SI 221

41

## analyse en composantes principales



L. Likforman - Telecom ParisTech - SI 221

42

## apprentissage

### ■ l'apprentissage consiste à:

- estimer des distributions de probabilités
- ou déterminer des frontières entre classes
- ou sélectionner/construire des prototypes
- ou construire des grammaires

### ■ 2 types d'apprentissage

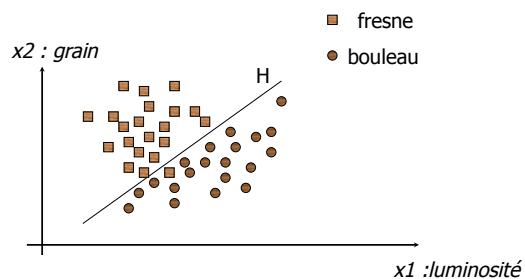
- supervisé :
  - formes + classes associées ( $x_i, c_i$ )  $i=1, N$
- non supervisé
  - formes + nbre de classes

L. Likforman - Telecom ParisTech - SI 221

43

## approche statistique (supervisée)

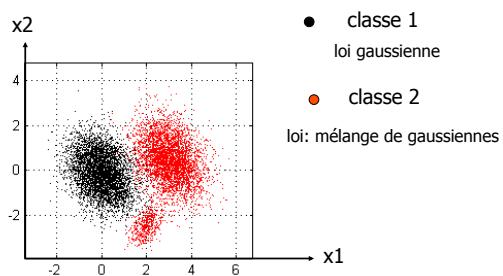
trouver une frontière dans l'espace de représentation  
(frontière explicite ou implicite)



L. Likforman - Telecom ParisTech - SI 221

44

## plus difficile ....

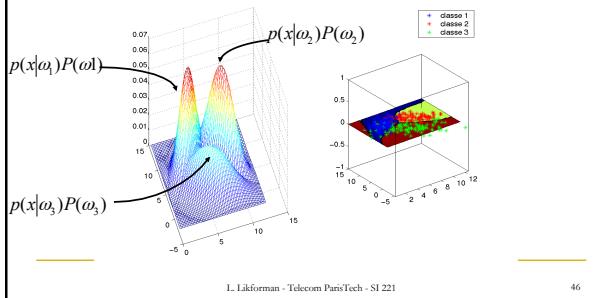


L. Likforman - Telecom ParisTech - SI 221

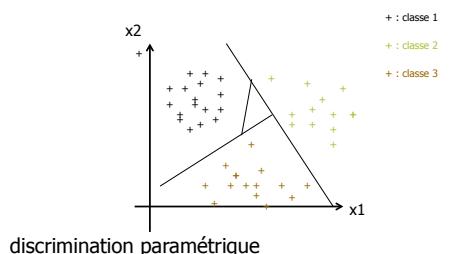
45

## approche statistique : décision Bayésienne

- approche optimale : connaissance des distributions sous jacentes



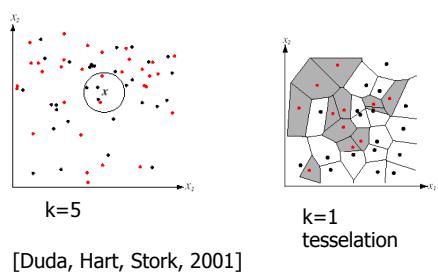
## approche statistique : séparation par des surfaces



L. Likforman - Telecom ParisTech - SI 221

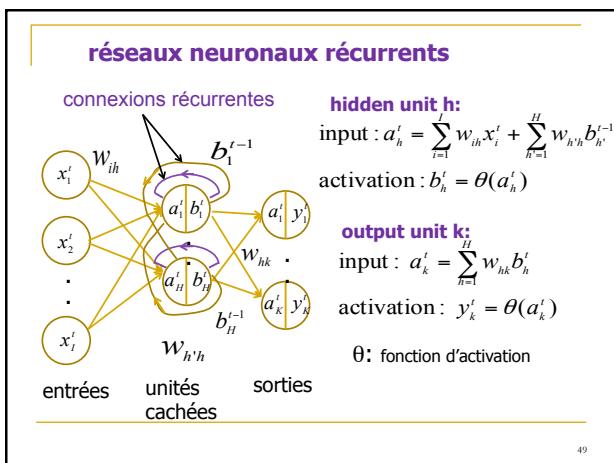
47

## approche statistique : k plus proches voisins



L. Likforman - Telecom ParisTech - SI 221

48



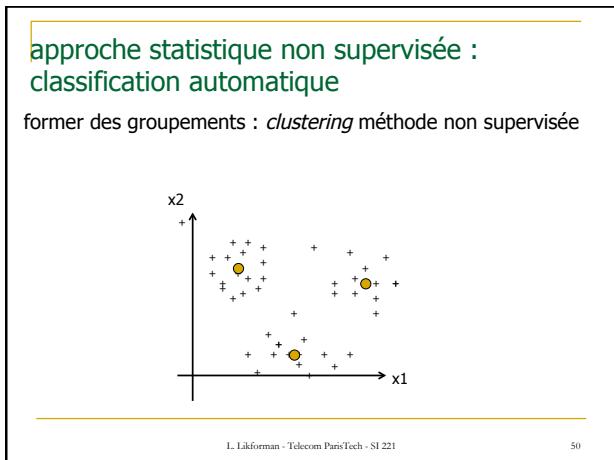

---

---

---

---

---



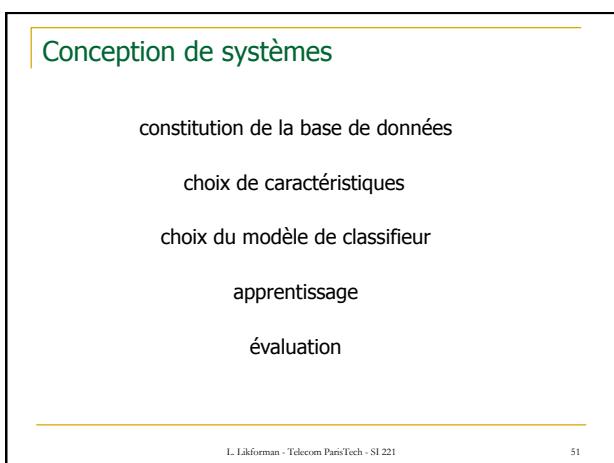

---

---

---

---

---




---

---

---

---

---

**Bases de données**

- UCI Repository of machine databases (307 bases)  
<http://www.ics.uci.edu/~mlearn/MLRepository.html>



**Wisconsin Breast Cancer Databases**  
 Original database  
 Donated by Olvi Mangasarian  
 Located in breast-cancer-wisconsin sub-directory, filenames root: breast-cancer-wisconsin  
 Currently contains 699 instances  
 2 classes (malignant and benign)  
 9 integer-valued attributes

L. Likforman - Telecom ParisTech - SI 221      52



**■ Internet Advertisements**  
 From Nicholas Kushmerick (nick@ucd.ie)  
 This dataset represents a set of possible advertisements on Internet pages.  
 The features encode the geometry of the image (if available) as well as phrases occurring in the URL, the image's URL and alt text, the anchor text, and words occurring near the anchor text. The task is to predict whether an image is an advertisement ("ad") or not ("nonad").  
 Number of Instances: 3279 (2821 nonads, 458 ads)  
 Number of Attributes: 1558 (3 continuous; others binary)

**■ Pen-Based Recognition of Handwritten Digits**  
 From E. Alpaydin, Fevzi Alimoglu  
 10 classes  
 7494 training cases, 3498 test cases  
 16 attributes (All input attributes are integers 0..100)

L. Likforman - Telecom ParisTech - SI 221 - 2010      53

**apprentissage**

**■ l'apprentissage consiste à:**

- estimer des distributions de probabilités
- ou déterminer des frontières entre classes
- ou sélectionner/construire des prototypes
- ou construire des grammaires

**■ 2 types d'apprentissage**

- supervisé :
  - formes + classes associées ( $x_i, c_i$ )  $i=1, N$
- non supervisé
  - formes + nbre de classes

L. Likforman - Telecom ParisTech - SI 221      54

## métriques

- distance euclidienne

$$d(X, Y) = \sqrt{\sum_{k=1}^d |x_k - y_k|^2}$$

- distance L1

$$d(X, Y) = \sum_{k=1}^d |x_k - y_k|$$

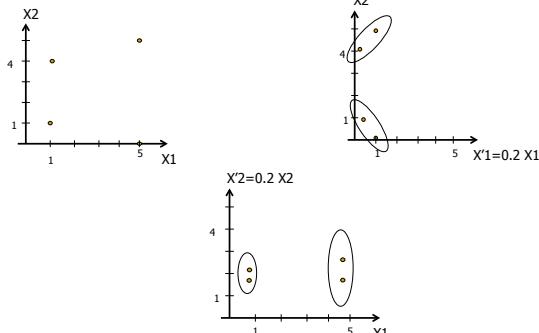
- distance de Chebychev  $L_\infty$

$$d(X, Y) = \max_k |x_k - y_k|$$

L. Likforman - Telecom ParisTech - SI 221

55

## effet d'échelle



L. Likforman - Telecom ParisTech - SI 221

56

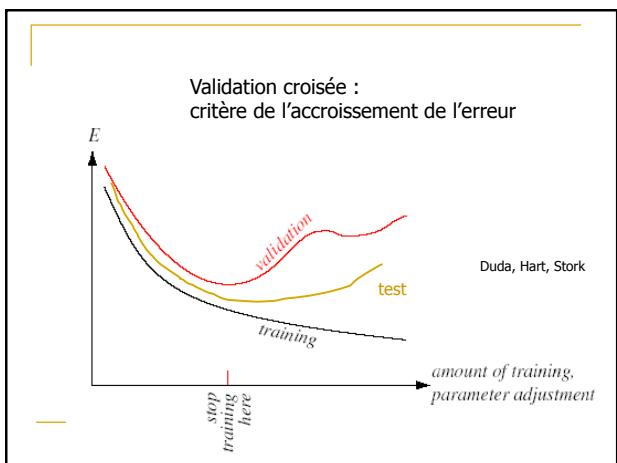
## apprentissage

- base de données découpée en 2 ou 3 parties
  - base d'entraînement (ou apprentissage) (*training*)
  - base de validation (si validation croisée)
  - base d'évaluation (ou test)

attention au sur-apprentissage (*over-fitting*)

L. Likforman - Telecom ParisTech - SI 221

57

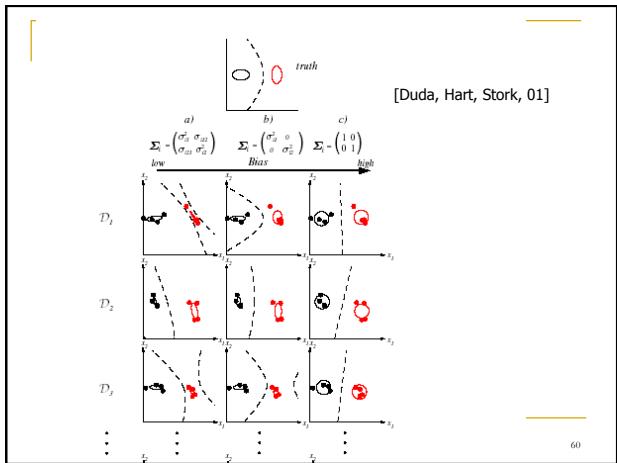


## dilemme biais-variance

- pour un autre ensemble d'apprentissage :
    - si frontière varie beaucoup -> variance élevée
    - si la frontière est proche de la frontière optimum -> biais faible
  - un classifieur de biais faible a une grande variance
  - privilégier une variance faible

J. Likforman - Telecom ParisTech - SI 221

59



## bibliographie

- S. Belongie, J. Malik, and J. Puzicha 2002 "Shape Matching and Object Recognition Using Shape Contexts". IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (24).
- D. Zhang, G. Lu; Review of shape representation and description techniques, Pattern Recognition 37 (2004)
- L. Likforman, E. Barney Smith, Reconnaissance des formes, théorie et pratique sous Matlab, Ellipses TechnoSup, 2013.
- R. Duda, P. Hart, D. Stork, Pattern Classification, Wiley, <http://www.dawsonera.com/abstract/9781118586006> + Login here +recherche auteur
- A. Cornuejols, L. Miclet, *Apprentissage artificiel : concepts et algorithmes*, Eyrolles, 2002.
- J. Wesley Hines, *matlab supplement to fuzzy and neural approaches in engineering*, Wiley, 1997.