

# IFT3395/6390

## Fondements de l'apprentissage machine

Premier cours:  
**INTRODUCTION**

Professeur: Pascal Vincent

# Au programme aujourd'hui

- Faire connaissance...
- Page Web du cours:  
Inscrivez-vous au cours IFT 3395/6390 sur le système Moodle: <http://moodle.iro.umontreal.ca>
- Présentation des objectifs, du plan de cours, des modalités d'évaluation, ...
- Présentation informelle du domaine de l'apprentissage automatique.

PREMIÈRE PARTIE

Plan de Cours  
et autres informations pratiques

# Les informations de plan de cours se trouvent sur:

• La page Web du cours:

<http://moodle.iro.umontreal.ca>

## Objectifs du cours

1

- **Technologie**: algorithmes d'apprentissage
  - pré-requis: algorithmique, structure de données, programmation, statistiques, algèbre linéaire
  - focaliser sur la **classification**
- **Philosophie**: la théorie de l'apprentissage
- **Applications**: ...

# Tonalité du cours

- Matériel pédagogique principal produit par **Balázs Kégl & Pascal Vincent**.
- Le cours nécessite d'être assez à l'aise en **mathématiques & informatique**.
- Le cours sera donné en Français, mais beaucoup de matériel est en Anglais (lectures, et matériel parfois utilisé comme support de cours).

## DEUXIÈME PARTIE

Présentation informelle du domaine  
de l'apprentissage automatique

# Au programme

- Place de l'apprentissage en **Intelligence Artificielle**.
- Les **disciplines fondatrices** de l'apprentissage.
- Les **domaines d'application** de l'apprentissage.
- Ex. de **types de problèmes** en apprentissage.
- La **représentation** des données.
- Ex. d'un **algorithme simple**.

David is 11 years old.  
 He weighs 60 pounds.  
 He is 4 feet, 6 inches tall.  
 He has brown hair.  
 His love is real.  
 But he is not.



A STEVEN SPIELBERG FILM  
**ARTIFICIAL INTELLIGENCE**

WARNER BROS. PICTURES and DREAMWORKS PICTURES  
 An AMBLIN/STANLEY KUBRICK Production A STEVEN SPIELBERG Film **A.I. ARTIFICIAL INTELLIGENCE** HALEY JOEL OSMENT  
 JUDE LAW FRANCIS O'CONNOR BRENDAN GLEESON and WILLIAM HURT *Robot Characters Designed by STAN WINSTON STUDIO*  
 Special Visual Effects & Animation by INDUSTRIAL LIGHT & MAGIC *Composer BOB RINGWOOD Music by JOHN WILLIAMS*  
 Film Editor MICHAEL KAHN, A.C.E. Production Designer RICK CARTER Director of Photography JANUSZ KAMINSKI, A.S.C.  
 Executive Producers JAN HARLAN WALTER F. PARKES Story by STEVEN SPIELBERG Based on a Screenplay by IAN WATSON  
 Directed by STEVEN SPIELBERG Produced by BRIAN ALDISS Produced by KATHLEEN KENNEDY STEVEN SPIELBERG BONNIE CURTIS  
**DREAMWORKS PICTURES** **WARNER BROS. PICTURES**  
 Directed by STEVEN SPIELBERG  
**SUMMER 2001**  
 A.C.L. Keyword: A.I. www.AImovie.com



HALEY JOEL OSMENT

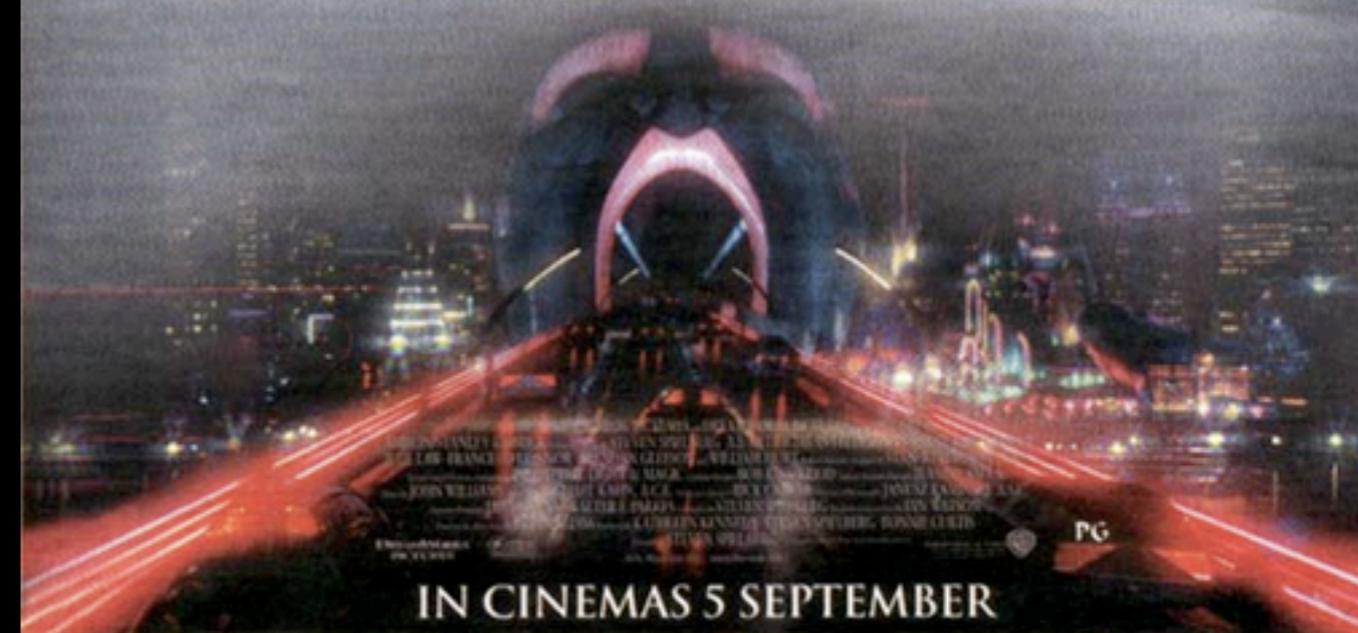
JUDE LAW

JOURNEY TO A WORLD WHERE ROBOTS DREAM AND DESIRE.

**A.I.**

A STEVEN SPIELBERG FILM

ARTIFICIAL INTELLIGENCE



IN CINEMAS 5 SEPTEMBER

PG



# L'Intelligence Artificielle, 2 approches historiques

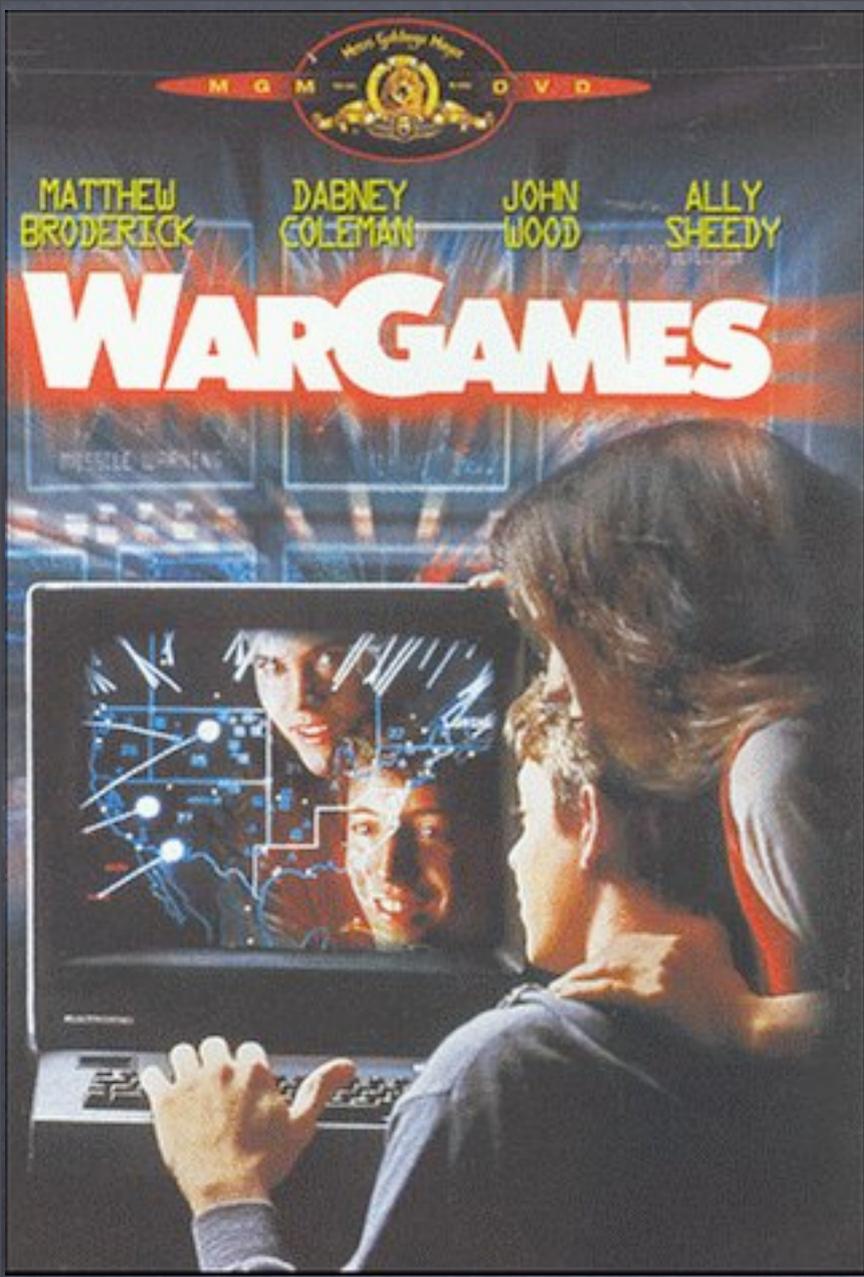
## I.A. "classique"

- Symbolique
- Systèmes à base de règles "logiques" construites à la main.  
Ex: systèmes experts.
- Succès moderne: Deep Blue bat Kasparov aux échecs en 1996.

## I.A. "connexionniste"

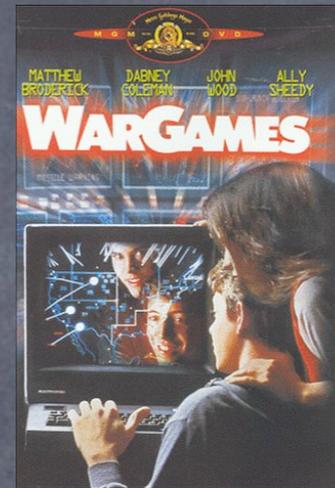
- Sub-symbolique
- Inspirée du fonctionnement du cerveau (réseaux neuronaux)
- L'**apprentissage** est très tôt au coeur de cette approche.
- "Elephants don't play chess"  
(R. Brooks, 1990)

1983



# De la science-fiction à la réalité...

- 1983: dans **WarGames**, un ordinateur apprend en jouant contre lui-même à **tic-tac-toe** et "**global thermonuclear war**".
- 1995: TD-gammon, un réseau neuronal entraîné en jouant **200 000 parties de backgammon** contre lui-même, joue à un niveau équivalent aux meilleurs joueurs mondiaux (Tesauro 1995).



Conclusion:  
THE ONLY  
WINNING  
MOVE IS NOT  
TO PLAY

# Vision de l'intelligence artificielle en 1957

(Rosenblatt, "Perceptron")

INFORMATIQUE

*Intelligence Artificielle*

Classique  
(symbolique)

Connexioniste  
(réseaux neuronaux)

NEURO-SCIENCES

# Place de l'apprentissage en I.A. en 2009

● L'I.A. "connexionniste" a mûri, s'est mathématisée, a engendré l'**apprentissage statistique** dont les réseaux neuronaux ne sont plus qu'une sous-partie.

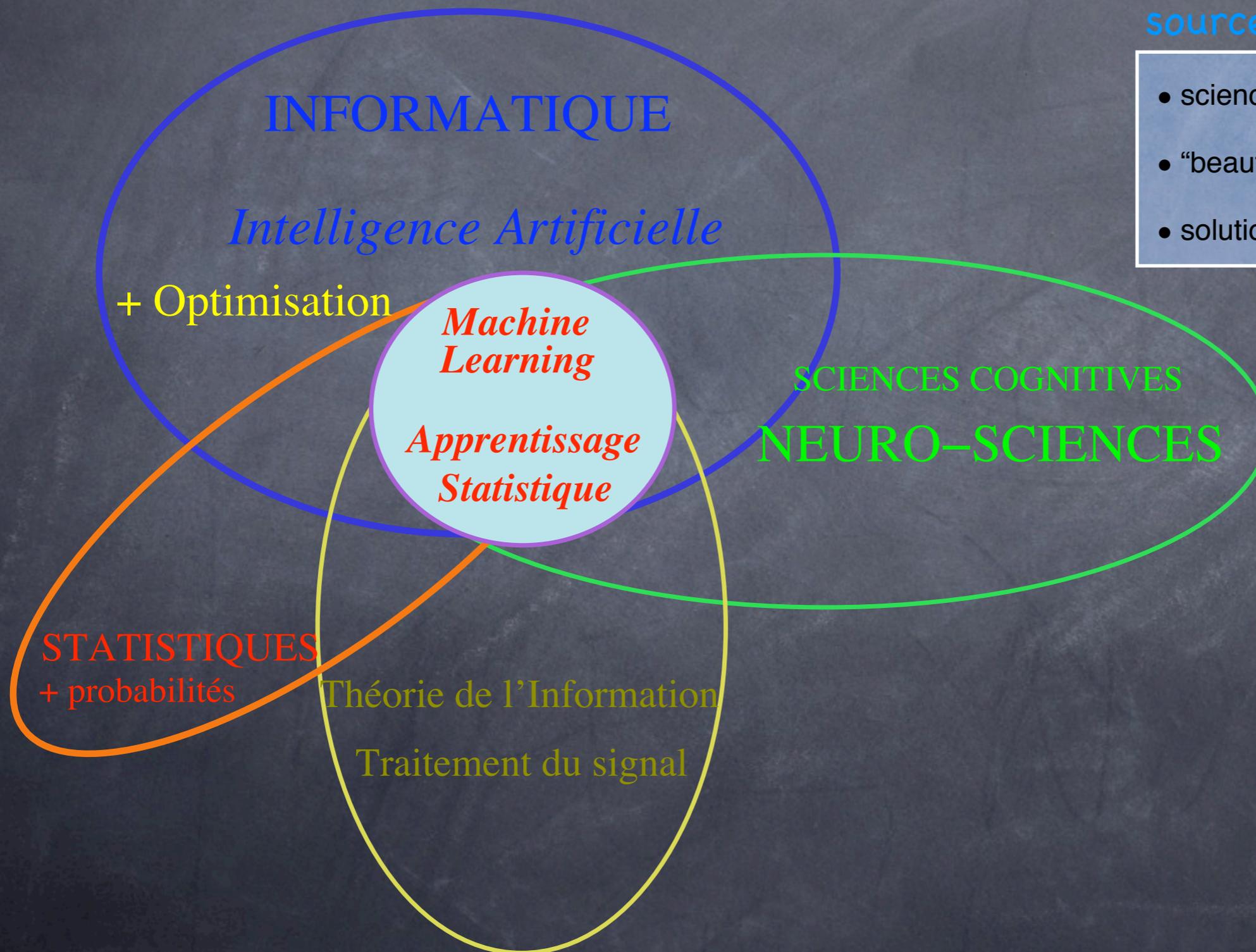
● L'I.A. classique, en intégrant l'incertain, a engendré les **modèles graphiques probabilistes** (réseaux Bayésiens), dont les paramètres peuvent être appris.

● Le rôle fondamental de l'**apprentissage** et d'une **approche probabiliste** est largement reconnu.

**apprentissage**

**statistique**

# Vision 2009 des disciplines fondatrices



sources d'inspiration:

- sciences cognitives/neurosciences
- "beauté mathématique"
- solutions d'ingénierie

# Apprendre à partir d'exemples !



“cheval”



“cheval”



“cheval”

Principe beaucoup plus général que d'écrire à la main, en partant de zéro, un algorithme pour reconnaître un cheval..

Vous savez programmer: comment feriez-vous ?

# Algorithmes “classiques” vs. d’apprentissage

- Approche classique

- description formelle des contraintes de l’entrée et de la sortie souhaitée
- compréhension du problème computationnel
- design d’une solution algorithmique basée sur ces connaissances

- Problèmes

- Connaissances incomplètes
- Algorithme trop coûteux

# Algorithmes “classiques” vs. d’apprentissage

- Approche d’apprentissage

- données (exemples) de forme (entrée, sortie)
- compréhension **partielle** du problème: connaissances a-priori
- apprendre: **chercher dans un ensemble** de fonctions

# Ex: la reconnaissance de caractères

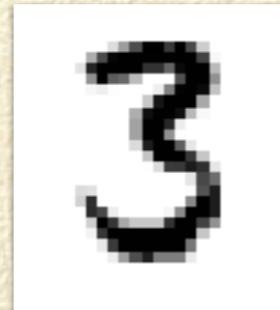
Ensemble d'apprentissage

2

3



Point de test



2 ou 3 ?

Apprendre n'est pas  
simplement  
*mémoriser...*

C'est être capable  
de *généraliser!*

# Apprentissage

- Une caractéristique essentielle de l'**intelligence** naturelle
- Apprendre **par coeur** vs. l'apprentissage **inductif**
- Mot clé: **GÉNÉRALISATION**
- Situation d'apprentissage typique:
  - 1 On nous donne des **exemples** (des données)
  - 2 On nous présente un **nouveau cas** et il faut prendre une décision

# Catégories d'apprentissage automatique

- Classification
  - classifier le nouvel exemple
- Régression
  - faire une prédiction à partir du nouvel exemple
- Estimation de densité
  - dire si le nouvel exemple ressemble aux exemples déjà vus
- (Modèles graphiques)

# Applications

- Applications “**traditionnelles**”: reconnaissance des formes
  - écriture, parole, empreinte digitale. etc.
  - les **experts humains** les font bien
  - nombre de **données**, nombre d'**attributs**, nombre de **classes** modérés
  - **bruit/ambiguïté** modéré

# Applications

- Applications “modernes”:
  - forage de données, forage de texte, prédiction financière, ranger des “hits”(Google), analyse des expressions génétiques
  - les experts humains n'existent pas
  - nombre de données, nombre d'attributs, nombre de classes énormes
  - bruit/ambiguïté augmenté

# Applications

- Reconnaissance des formes
  - écriture
  - parole
  - empreinte digitale
- Forage de texte
  - Google
  - classification de textes

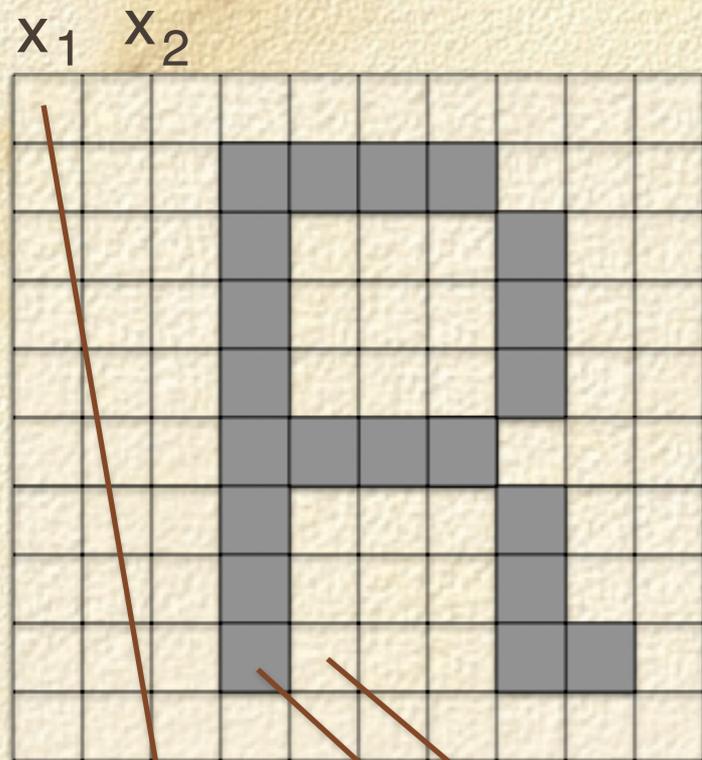
# Applications

- Forage de données
  - marketing direct
  - prédiction des primes
- Bioinformatique
  - prédiction du risque de cancer
  - détection de cancer
  - découverte de médicament

# Applications

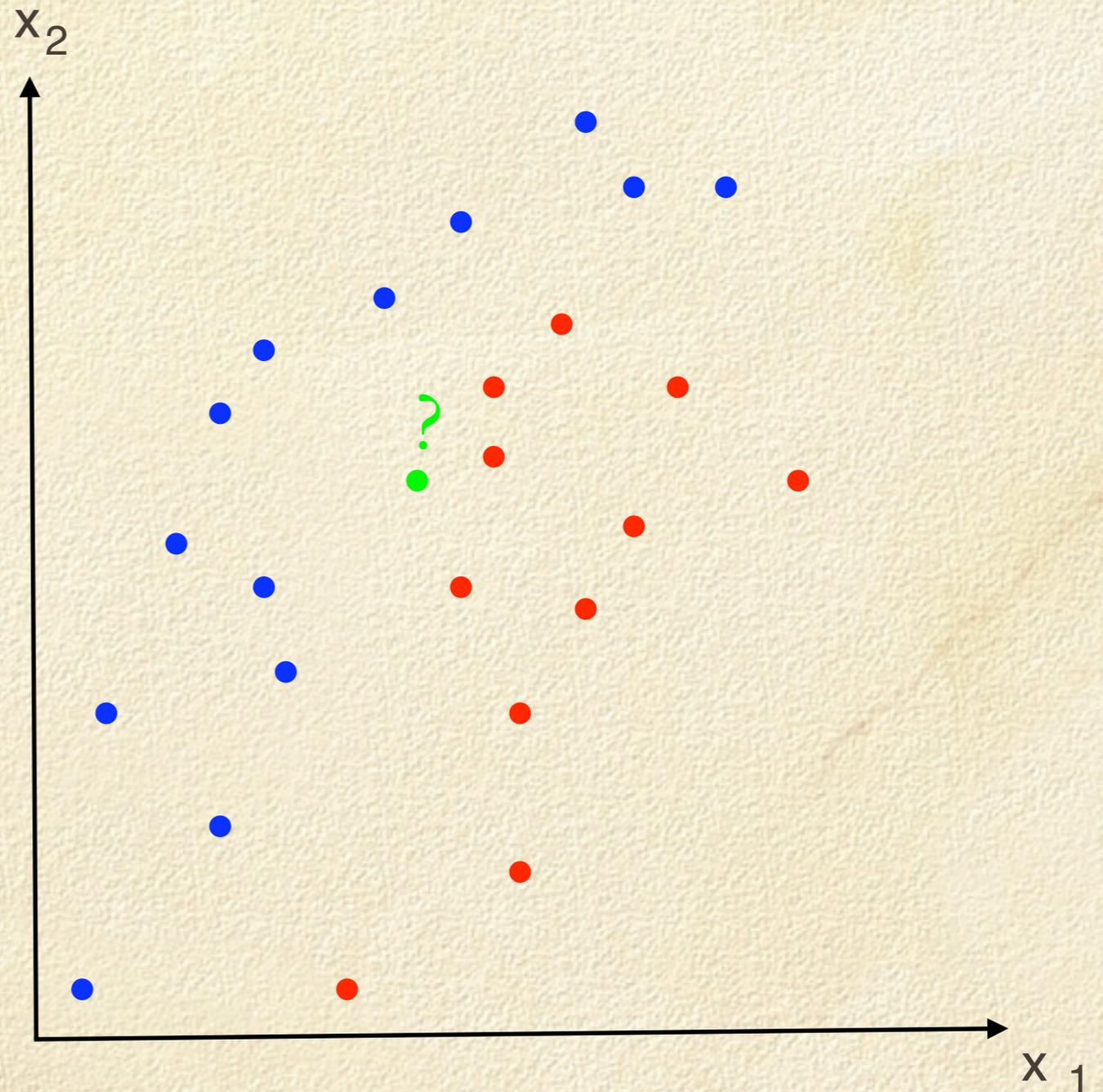
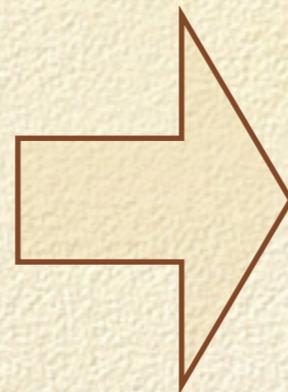
- Traitement de langage
  - prédiction du prochain mot
  - désambiguïsation du sens
  - prédiction du POS (*Part-Of-Speech*)
- Génie logiciel
  - prédiction de la stabilité

# Représentation des données



$$x = (0, 0, \dots, 140, 0, \dots)$$

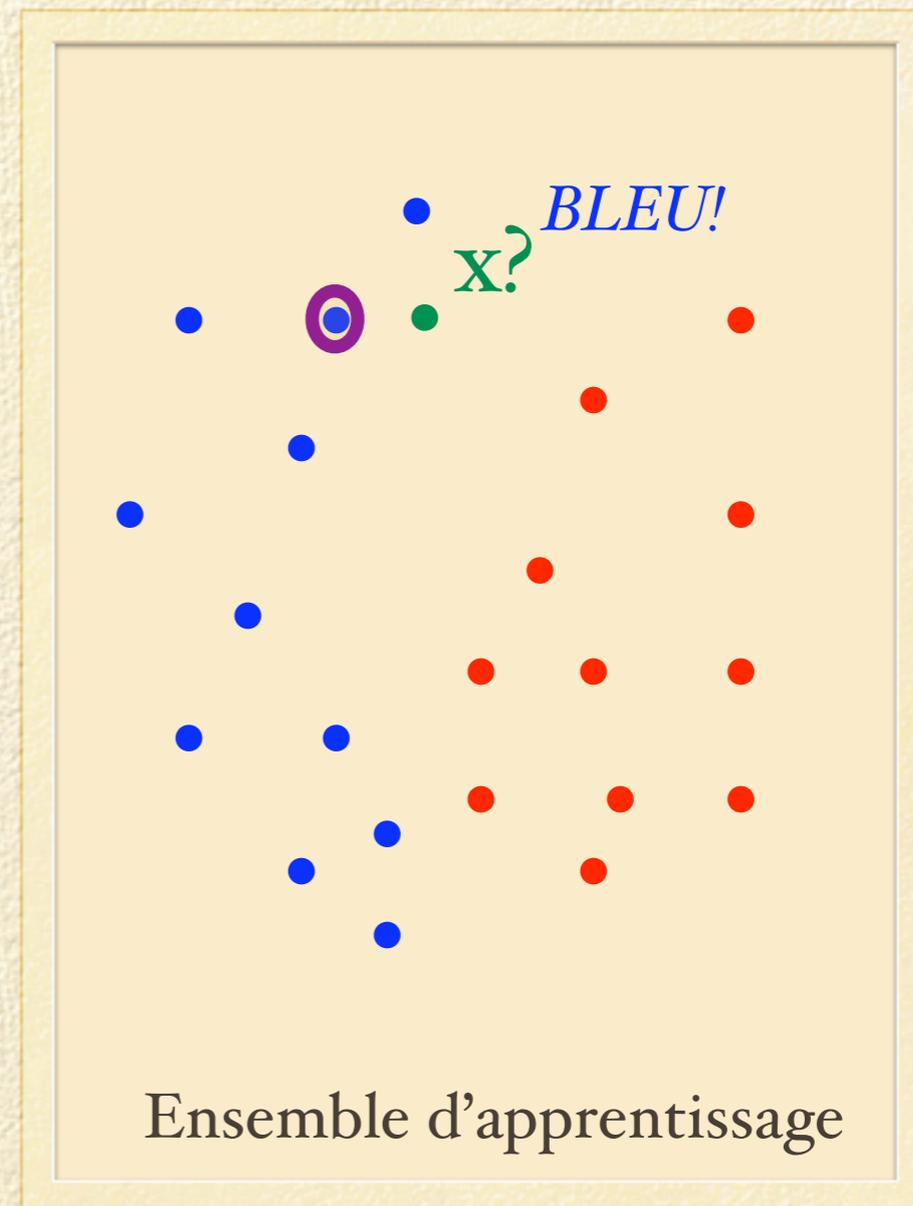
$x$  vecteur de  $\mathbb{R}^n$



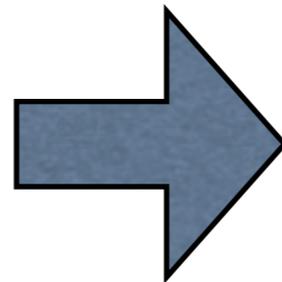
# L'algorithme du plus proches voisins

Pour un point test  $x$ :

- On trouve **le plus proche voisin** de  $x$  parmi l'ensemble d'apprentissage selon une certaine mesure de distance (ex: distance Euclidienne).
- On associe à  $x$  la classe de ce plus proche voisin.



# La notion de niveau de représentation



raw input vector representation:

$$\mathcal{X} = \begin{bmatrix} 23 & 19 & 20 & \dots & 18 \end{bmatrix}$$

$x_1$     $x_2$     $x_3$     $x_n$

slightly higher level representation

... etc ...

very high level representation:

CAT   JUMPING