

Datasciences & IA

Christophe Rodrigues

Rappel – les différents types de recherche

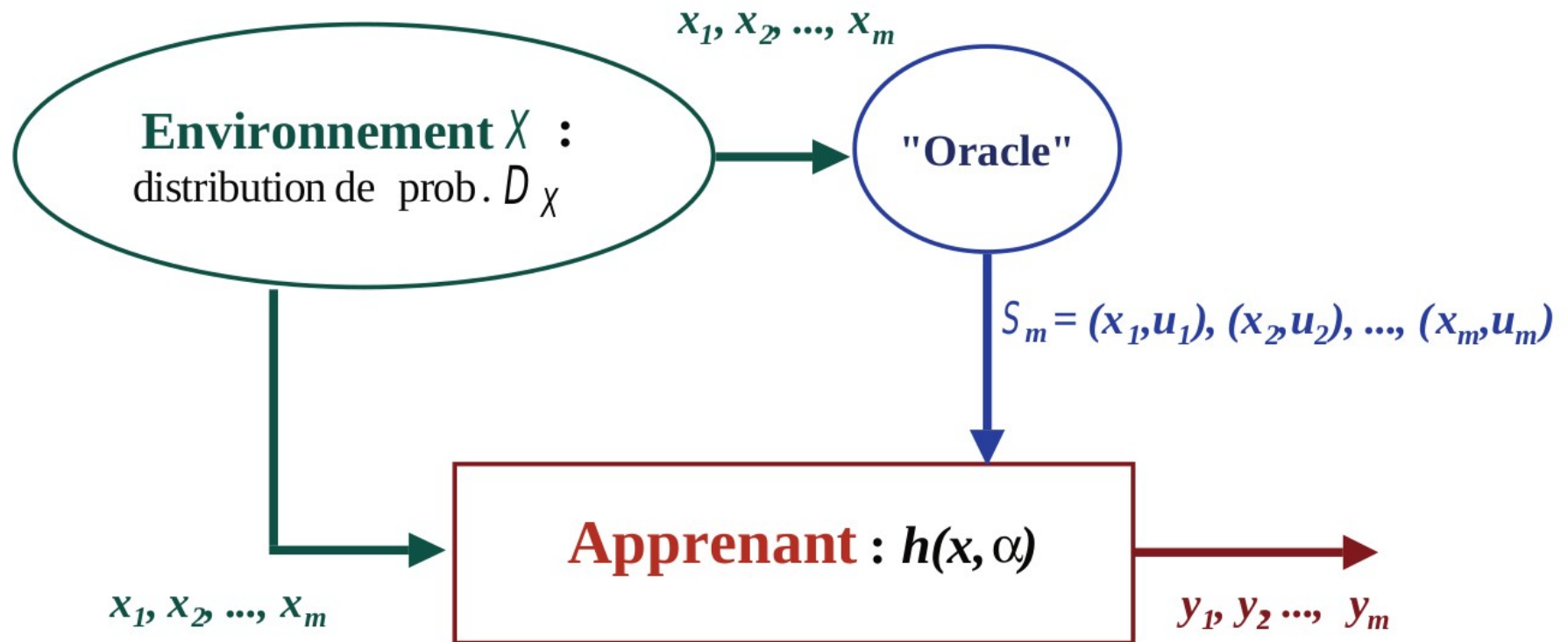
- **Algorithme génétique**
 - **Apriori – A***
 - **Programmation par contrainte**
 - **En présence d'un adversaire**
-
- **Recherche d'hypothèses dans un espace d'exemples :**

Apprentissage Automatique (Machine Learning)

Plan du cours

- **Définition de l'apprentissage**
- **Evaluation de l'apprentissage**
- **Exemple de méthode supervisée (kNN)**
- **Exemple de méthode non-supervisée (DBSCAN)**
- **Historique de l'apprentissage**

Apprentissage par induction

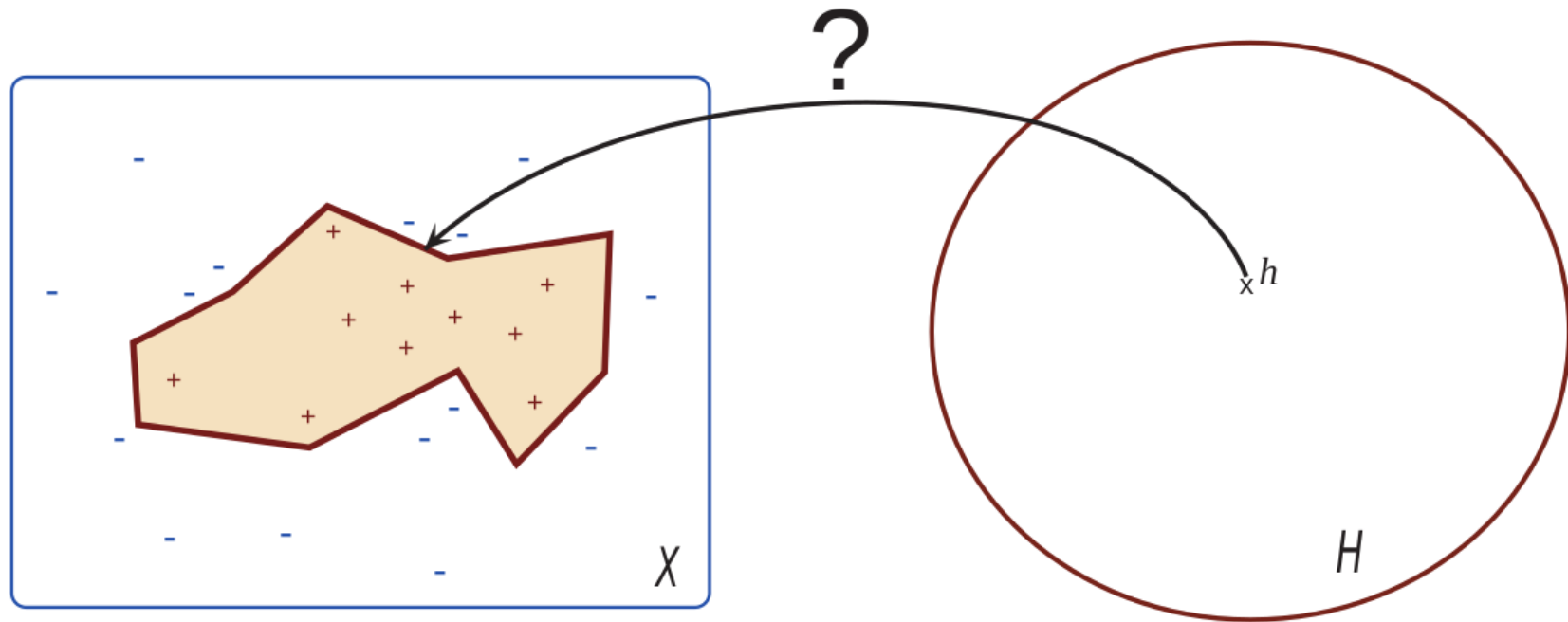


Définition de l'apprentissage

Espace des exemples : X

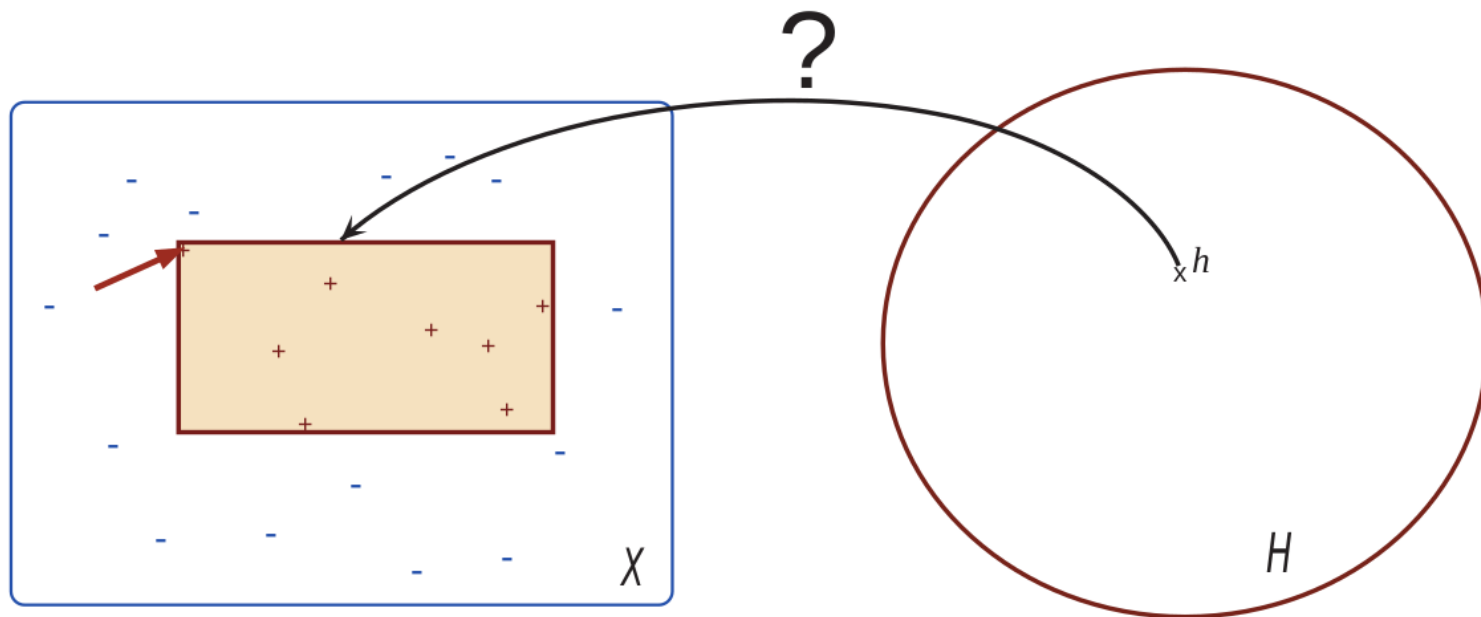
Espace des hypothèses : H

On Cherche une hypothèse h capable
d'expliquer au mieux les exemples



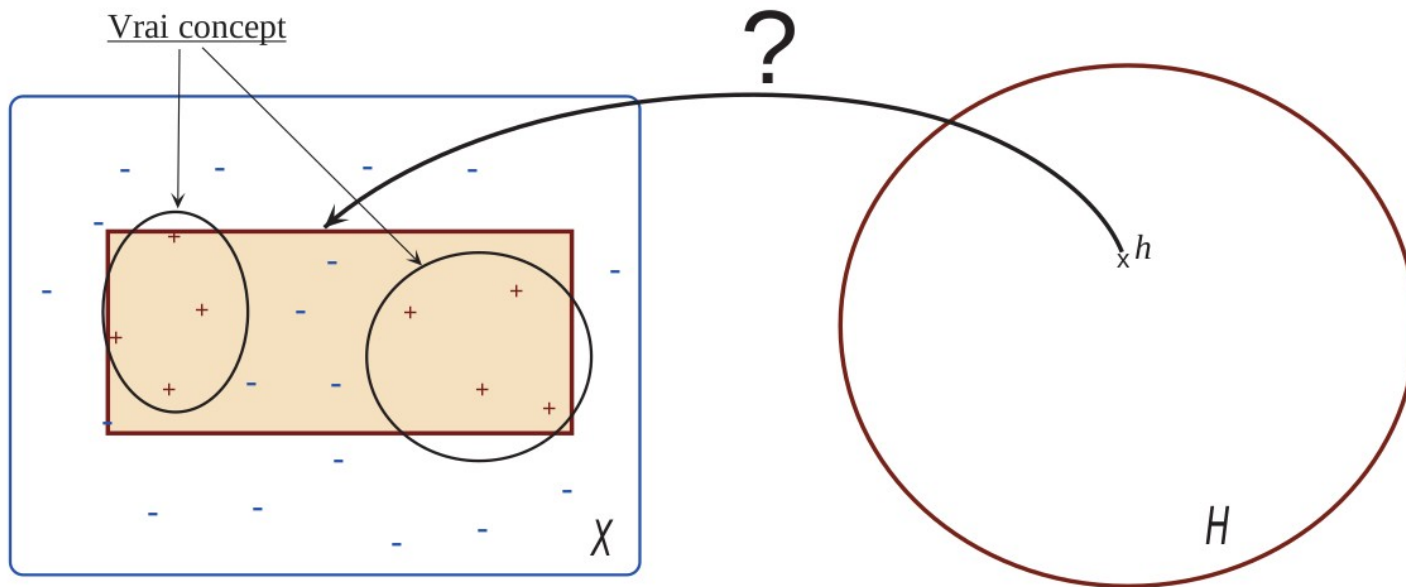
Choix d'un langage d'hypothèses

Exemple des hypothèses « rectangles » :
Définition d'un langage

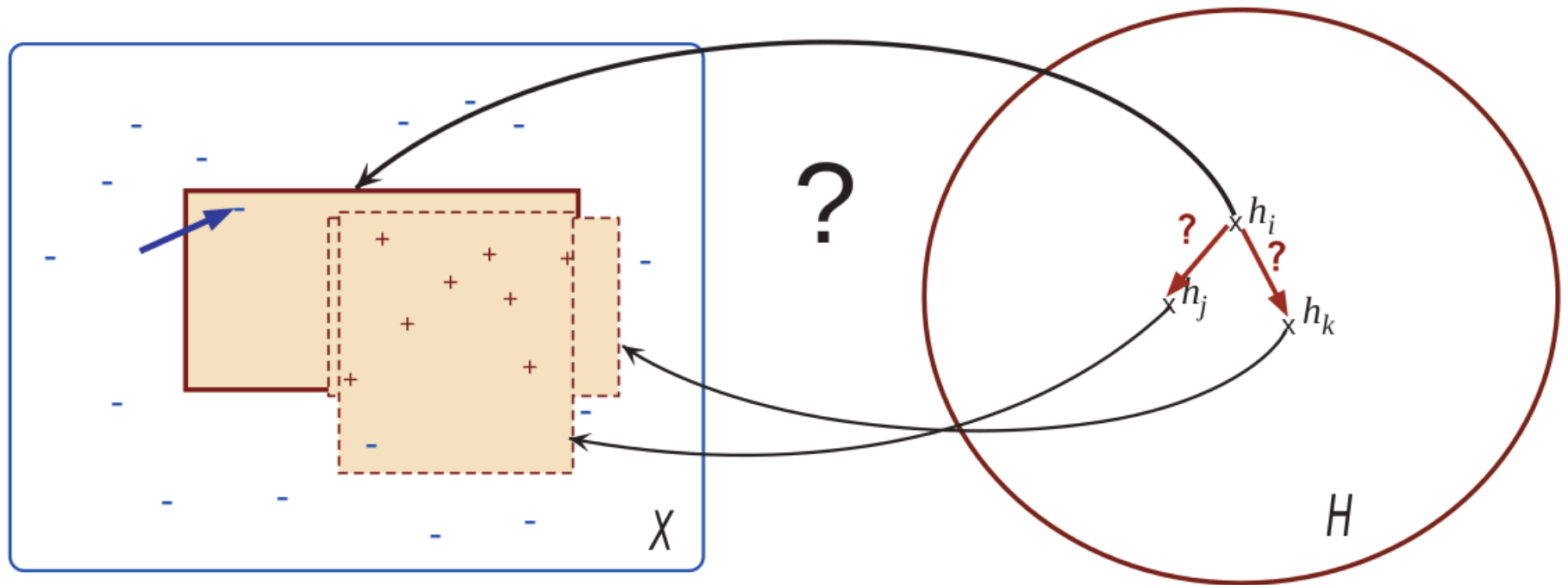


Expressivité d'un langage

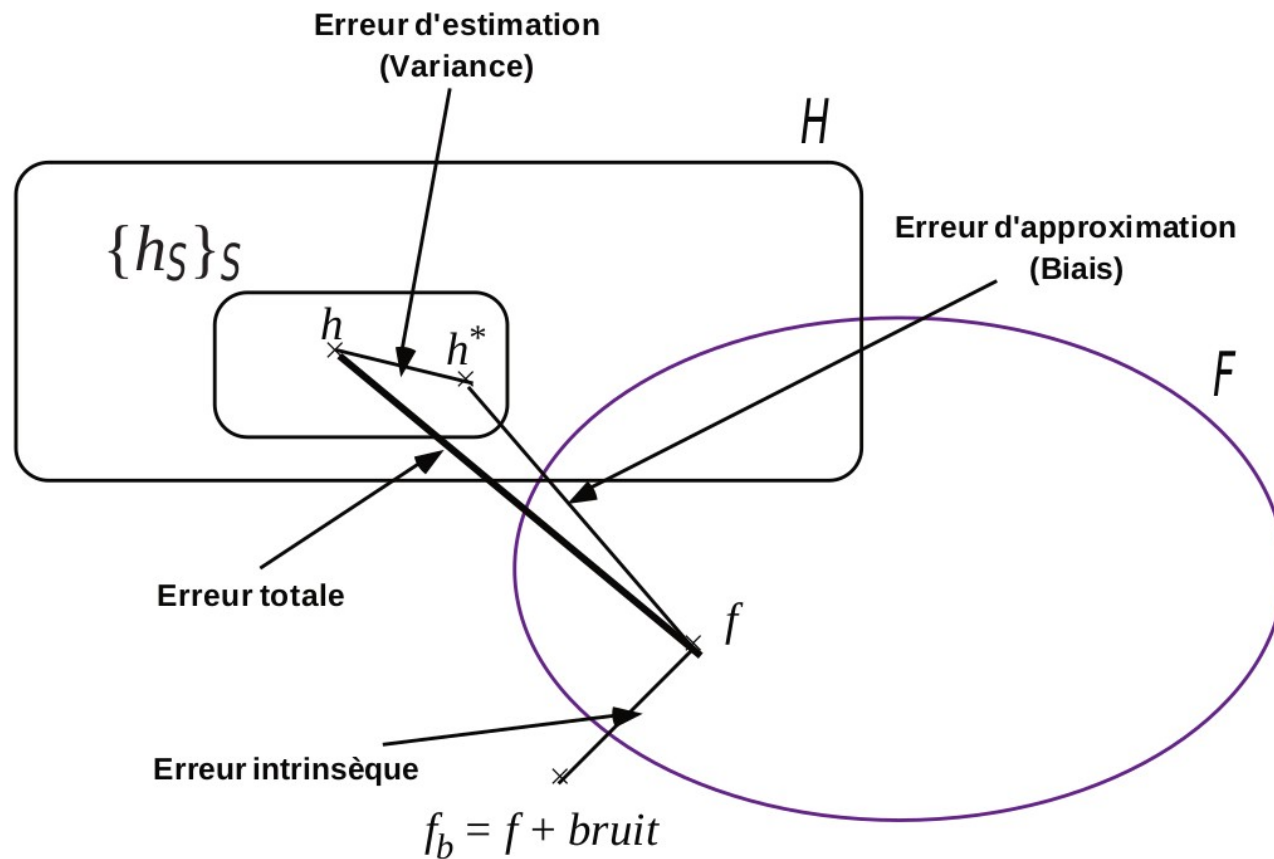
Impossible dans ce cas avec une
hypothèse « rectangle » de ne pas faire
d'erreurs de prédiction



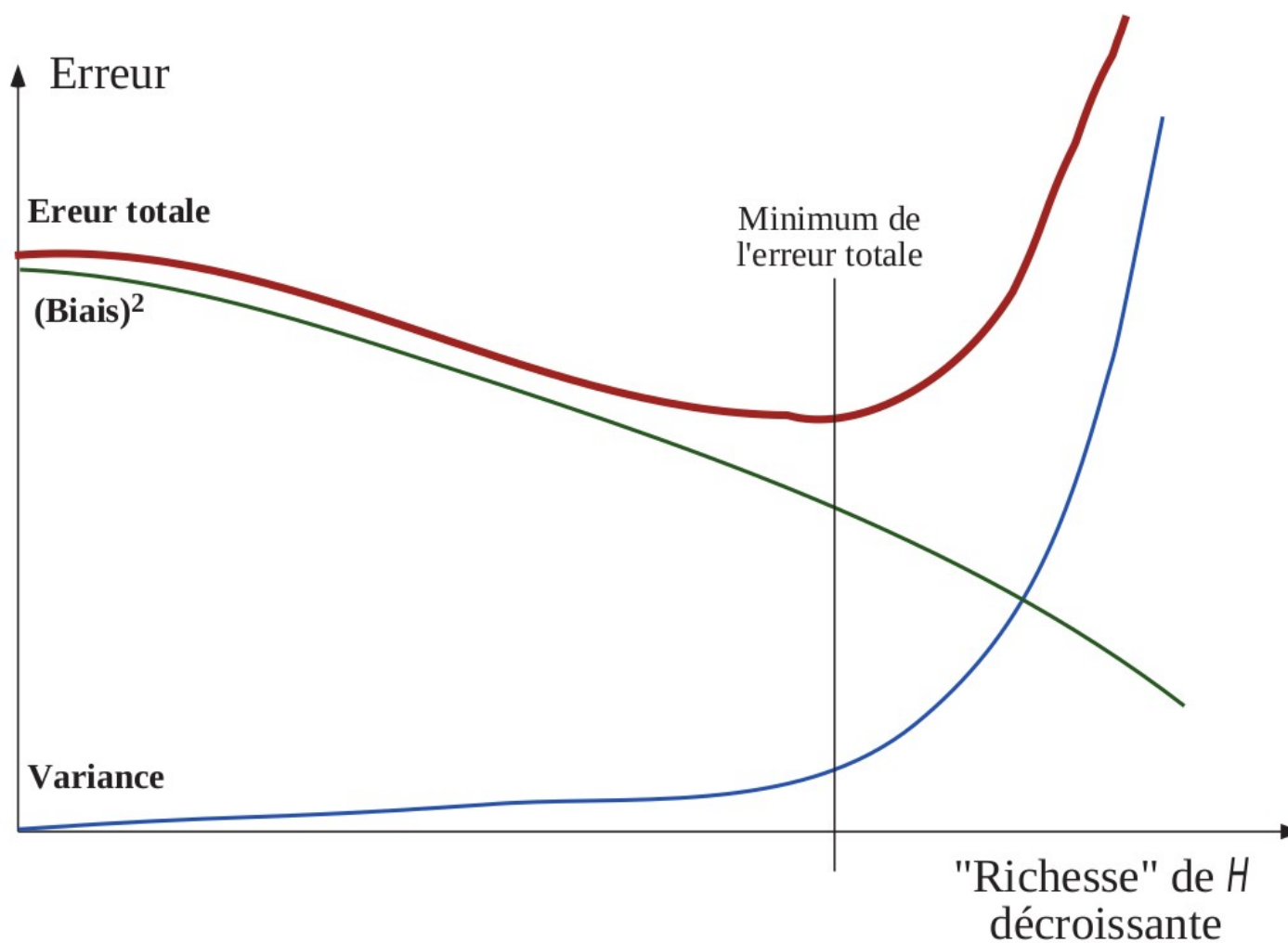
Exploration de l'espace des hypothèses



Les erreurs de l'apprentissage



Le compromis Biais/Variance



Apprentissage automatique (ML)

- **Apprentissage supervisé**

- Classification (prédire une catégorie)
- Régression (prédire une valeur continue)

- **Apprentissage non-supervisé**

- Regroupement (partitionner les exemples)

Evaluation de l'apprentissage

- **Comment tester si une hypothèse est assez générale et éviter le surapprentissage ?**
- **Séparation des données en deux groupes :**
 - Données d'apprentissage
 - Données d'évaluations
 - (voir trois) Données de validations
- **Répartition des données par tirage aléatoire**

Evaluation de l'apprentissage - Erreur

Dans le cas binaire de l'exemple :

Matrice de confusion

	'+'	'-'
'+'	Vrais positifs	Faux positifs
'-'	Faux négatifs	Vrais négatifs

Généralisable à plusieurs classes

Dans le cas continu pour N exemples :

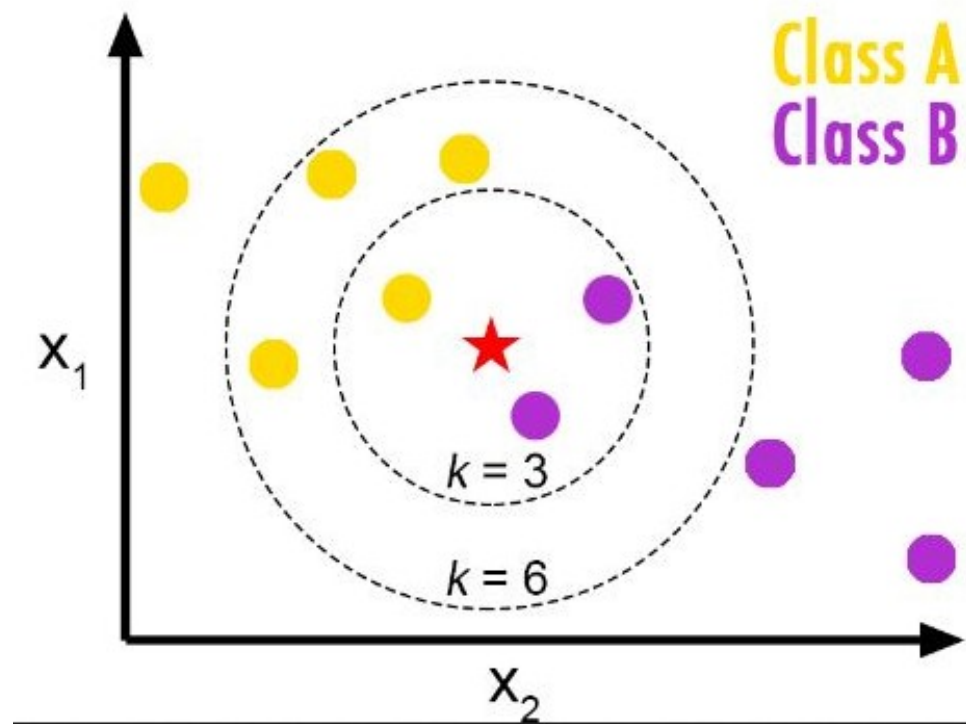
$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Exemple supervisé - k plus proches voisins

Repose sur la densité

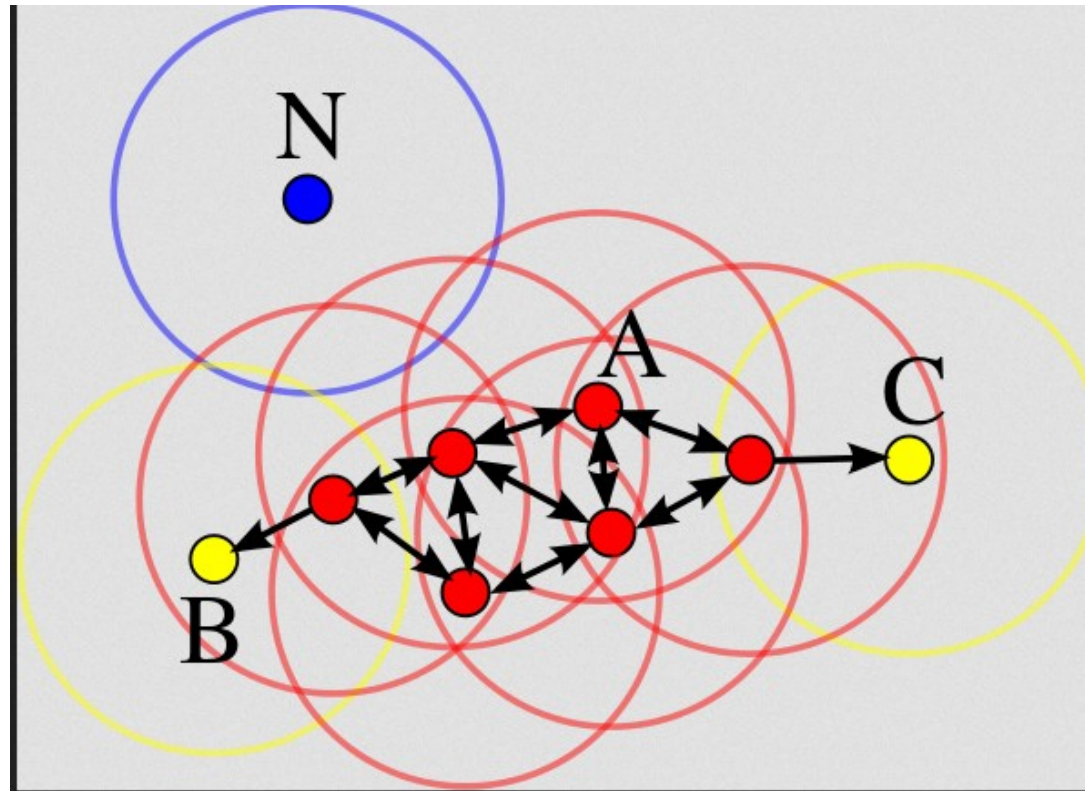
Généralisable à plusieurs dimensions, plusieurs classes ou en régression

Dépend fortement du paramètre k

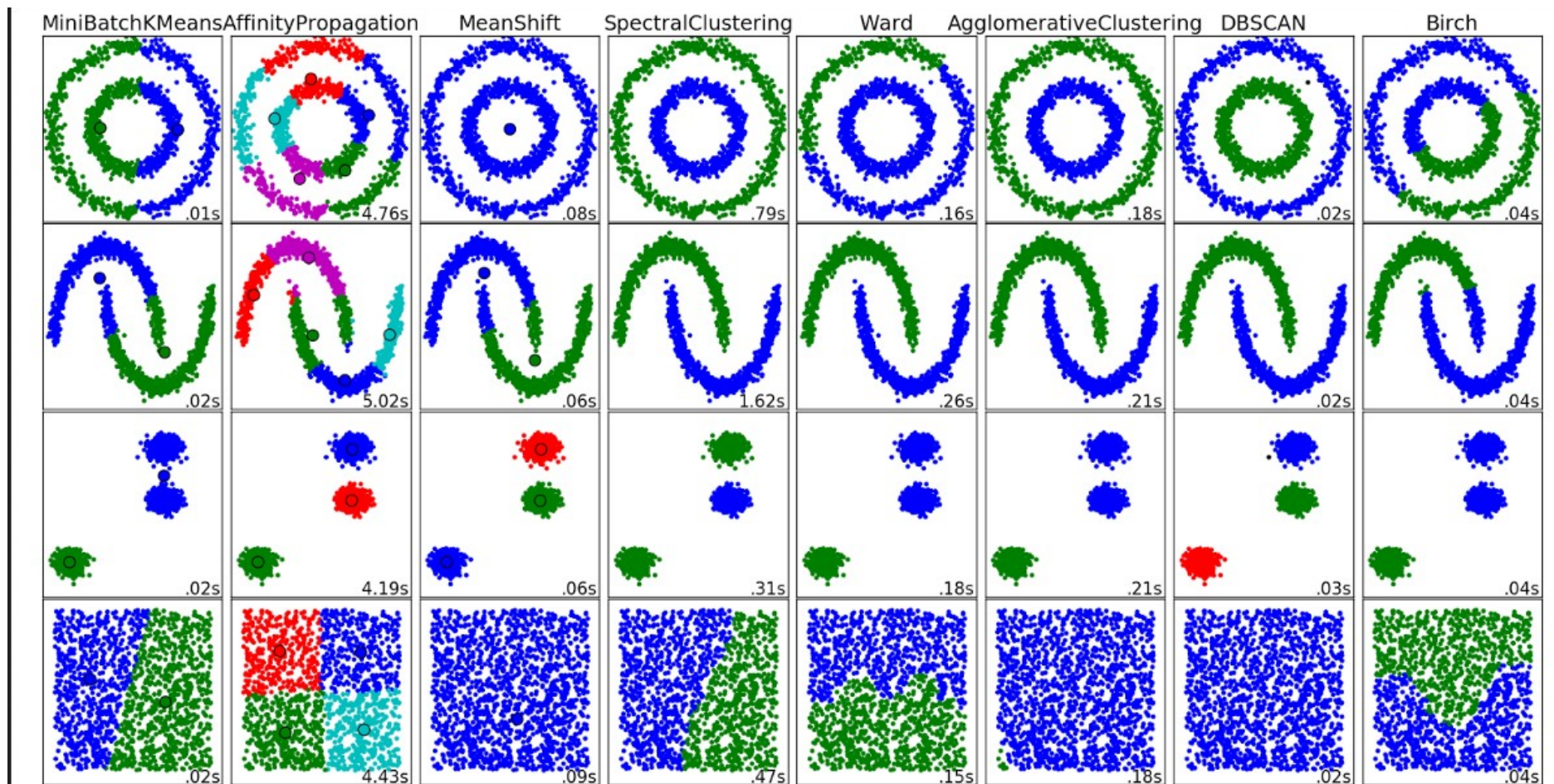


Exemple non-supervisé - DBSCAN

Deux paramètres : epsilon (taille du cercle) et MinPts (nombre minimal de points par cercle pour créer un cluster)



Comparaison de méthodes de clustering



Évolution de l'apprentissage automatique

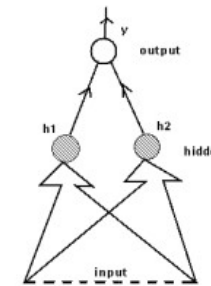
Discriminant
linéaire (Fisher)

Turing
(naissance de
l'informatique)

Expériences :
tortues cybernétiques



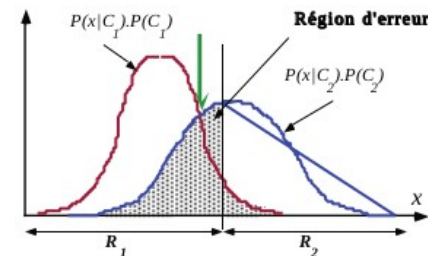
Perceptron



h1	h2	y
0	0	0
0	1	1
1	0	1
1	1	1

A corresponds to '1'

Reconnaissance des Formes :
Théorie de la décision bayésienne



20s

1936

50s

60s

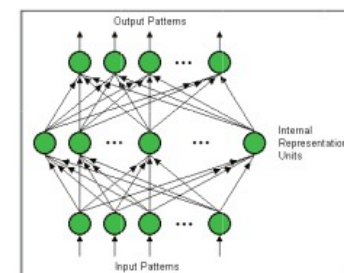
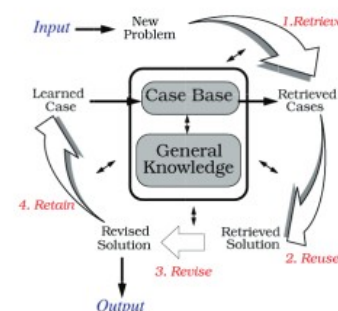
Évolution de l'apprentissage automatique

Apprentissage artificiel : une explosion

Systèmes dédiés à une tâche :
inspiration psychologique

Induction supervisée

Arbres de décision
Algorithmes génétiques
Explanation-Based Learning
Raisonnement par cas



ARCH

AM

META-DENDRAL

2^{ème} connexionnisme

1970

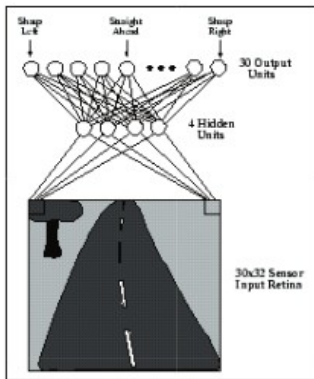
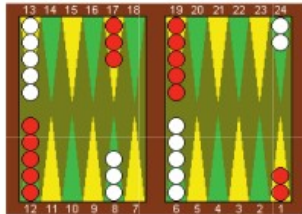
1976

1978

70s

80s

Évolution de l'apprentissage automatique



Apprentissage artificiel :
une théorisation
et une mise à l'épreuve



Théorie de Vapnik

Nouvelles méthodes :

- SVMs
- Boosting

Data mining
Text mining

