

Institut
québécois
d'intelligence
artificielle

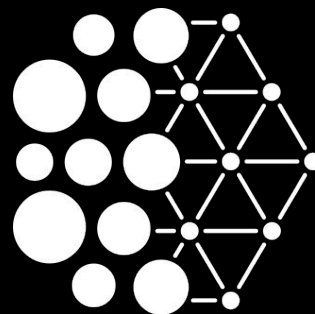
Mila

Université 
de Montréal



McGill

Institut
québécois
d'intelligence
artificielle



Mila

Bases de l'apprentissage automatique

Tristan Sylvain
tristan.sylvain@gmail.com

Apprentissage automatique

Champ de l'**intelligence artificielle** qui consiste à apprendre à un **système informatique** à réaliser une tâche à partir de **données**.

“On dit qu’un programme d’ordinateur apprend d’une expérience E par rapport à une classe de tâches T et une mesure de performance P , si sa performance sur la tâche T mesurée par P s’améliore avec l’expérience E .”
(Mitchell, 1997)

Quelles sont les différentes catégories d'apprentissage ?

Types d'apprentissage

- Supervisé
- Semi-supervisé
- Non-supervisé
- Par renforcement

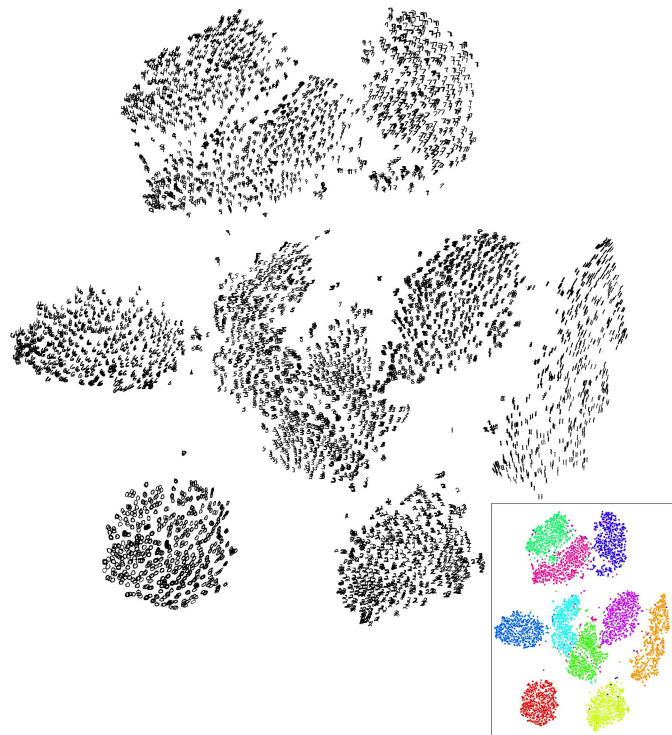
Apprentissage supervisé

- Chaque exemple associé à une ou plusieurs cibles
- Exemples: prédiction de la survie d'un patient, classification en radiologie.
- Tâche la plus classique.



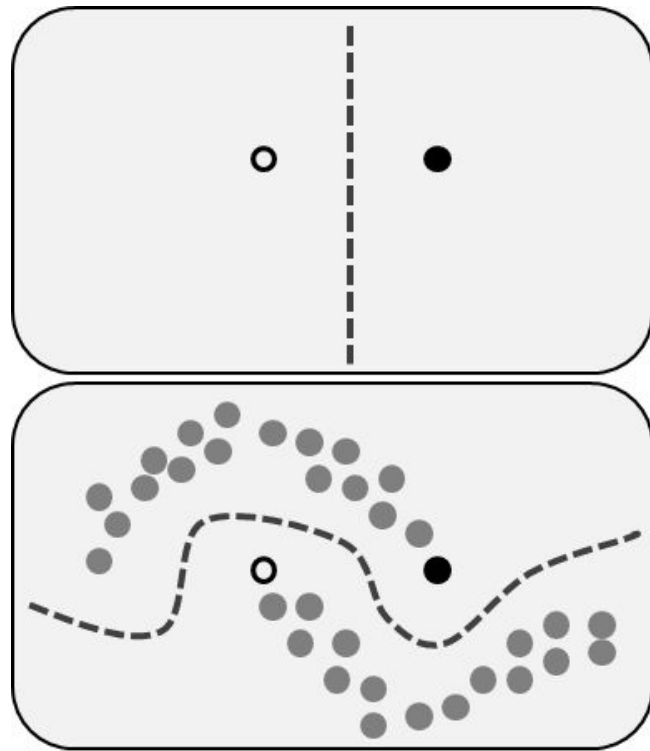
Apprentissage non-supervisé

- On ne dispose pas de cibles.
- On veut donc apprendre des caractéristiques du jeu de données, telles des grappes (clusters) de points.



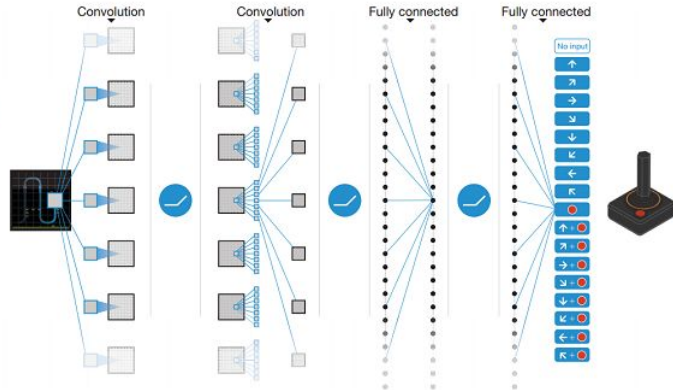
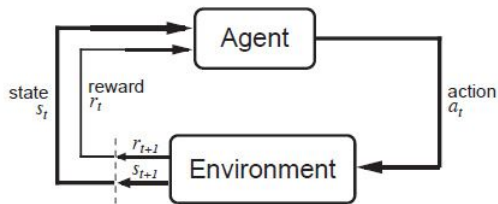
Apprentissage semi-supervisé

- Cas hybride des deux précédents: mélange de données avec et sans cible.
- Intéressant car survient dans beaucoup de cas en pratique.



Apprentissage par renforcement

- Le système effectue une série d'actions et reçoit une récompense.
- But: maximiser la récompense.



Quelles sont les tâches classiques?

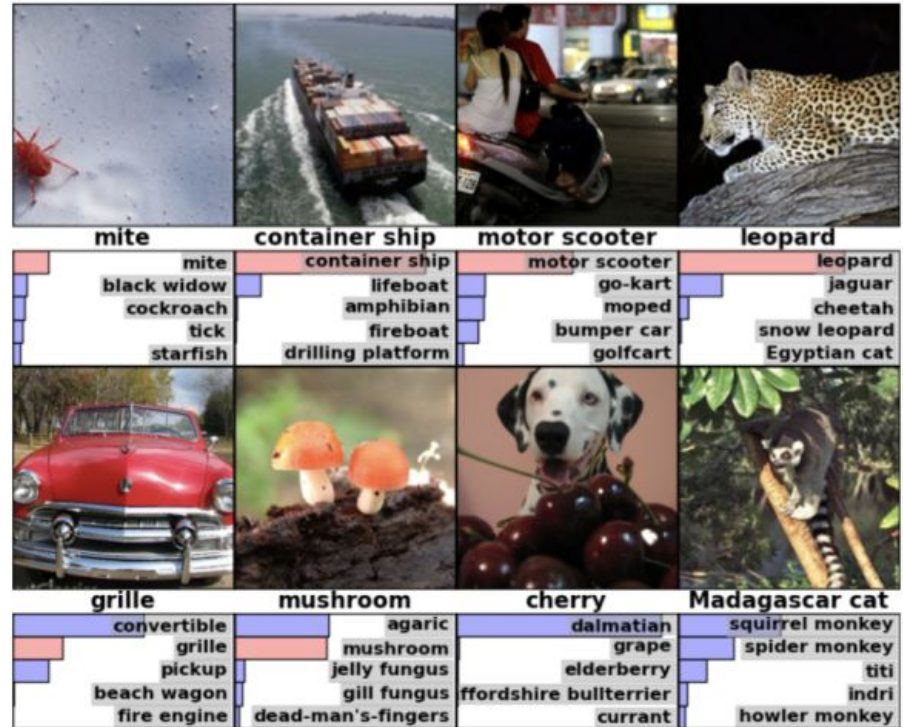
Régression

- On souhaite prédire une valeur continue.
- Exemples: series financières, temps de survie, age d'un patient à partir de l'image

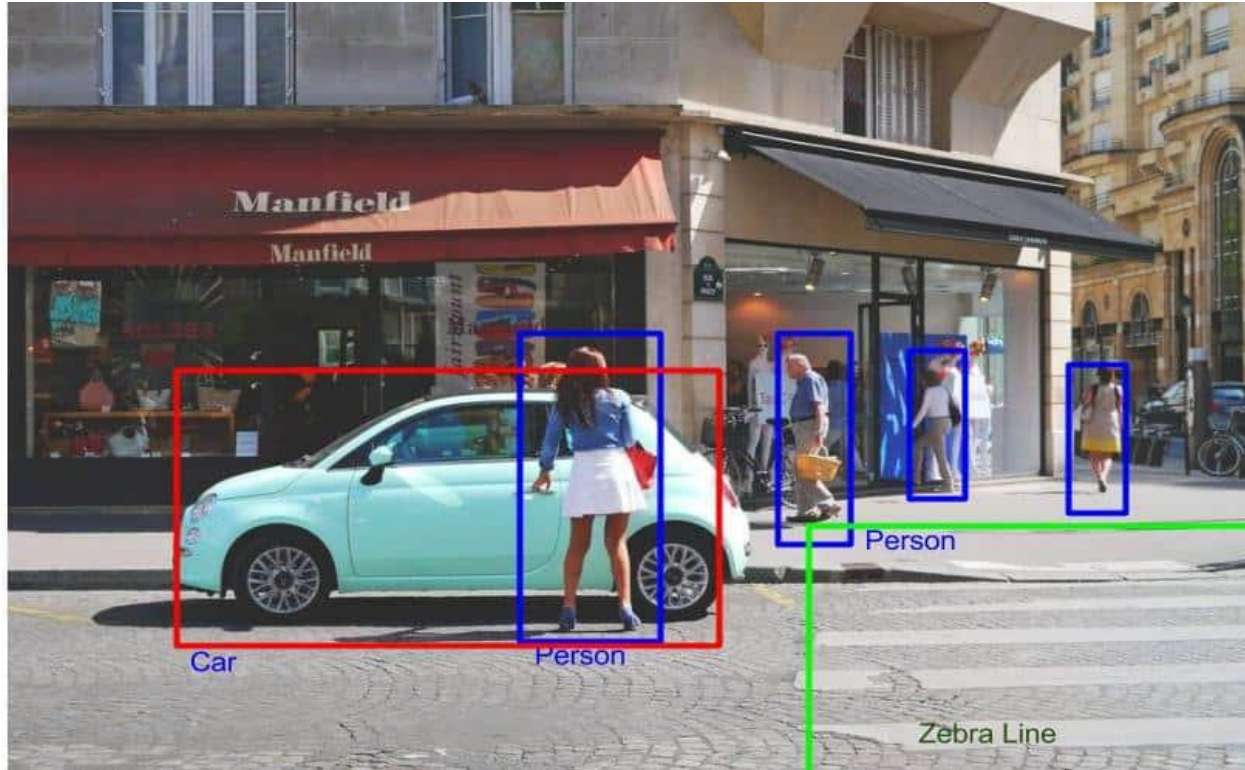


Classification

- On prédit une ou plusieurs classes.
- Multi-classe ou mono-classe
- Exemples: ImageNet, MNIST (chiffres), Classification de radios de patients, ...

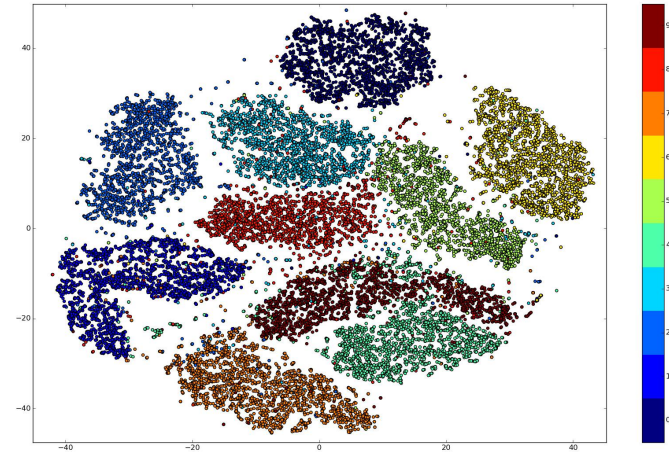


Reconnaissance d'objets



Clustering

- On veut faire apparaître des groupes de points.
- Tâche non-supervisée en general.



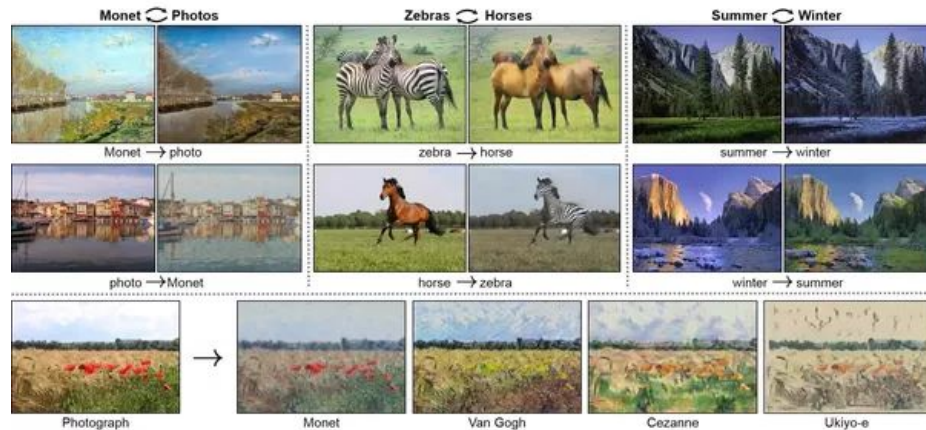
Génération

- On veut générer des échantillons proches de ceux vus pendant l'entraînement selon une certaine mesure
- Exemples: GANs, VAEs...



Adaptation de domaine

- On veut prendre des données dans un domaine, et les transposer dans un autre.
- Exemple: transformer une image de patient sain en patient malade.

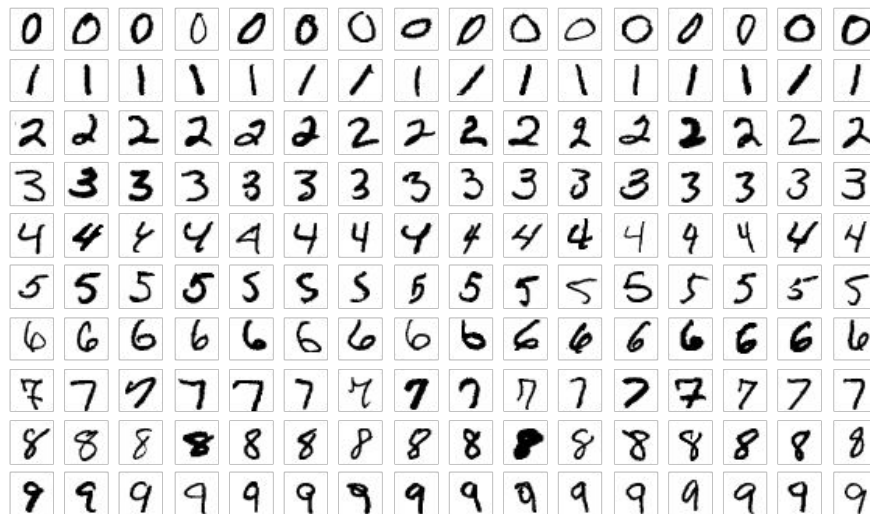


Comment faire en pratique ?

Données: nature et préparation

Base de données

- Contient des exemples, avec potentiellement une ou plusieurs cibles.
- Nature: images, texte, données multimodales, vidéos, ...



Représentation des données

- Image: ensemble de pixels.
- Texte: Bag of words, représentation par vecteur appris (Word2Vec, Glove...)



Exemple d'IRIS (1)

Variables $X \in \mathbb{R}^4$

Cible y

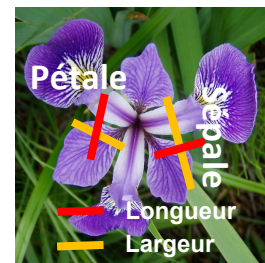
Exemples
(n = 150)

Longueur des sépales	Largeur des sépales	Longueur des pétales	Largeur des pétales	Cible
5,1	3,5	1,4	0,2	Setosa
4,9	3,0	1,4	0,2	Setosa
...				
6,4	3,5	4,5	1,2	Versicolor
...				
5,9	3,0	5,0	1,8	Virginica

Iris Setosa



Iris Versicolor



Iris Virginica



Exemple d'IRIS (2)

Représentation de l'entrée

	Représentation naturelle	Représentation à l'entrée du modèle			
Continue	5,2	<table border="1"><tr><td>5,2</td></tr></table>	5,2		
5,2					
Catégoriel	versicolor	<table border="1"><tr><td>0</td><td>0</td><td>1</td></tr></table> Vecteur One-hot	0	0	1
0	0	1			

Exemple d'IRIS (3)

Prétraitement des données

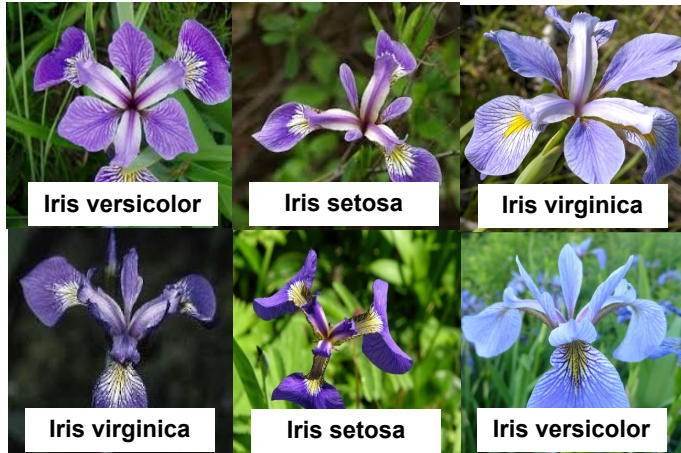
Normaliser les données X et les données y

❖ Standardisation $\tilde{x} = \frac{x - \mu}{\sigma}$

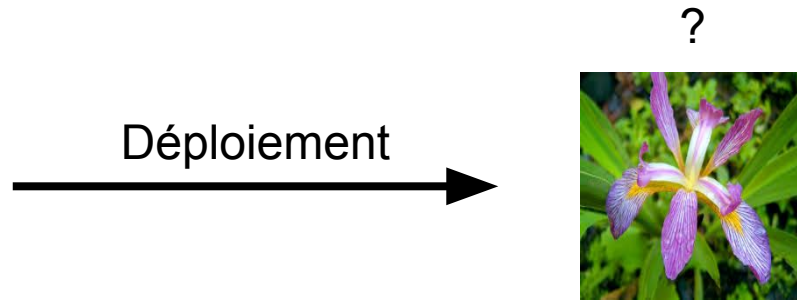
❖ Normalisation $\tilde{x} = \frac{x - x_{min}}{x_{max} - x_{min}}$

Exemple d'IRIS (4) Déploiement

Ensemble d'entraînement

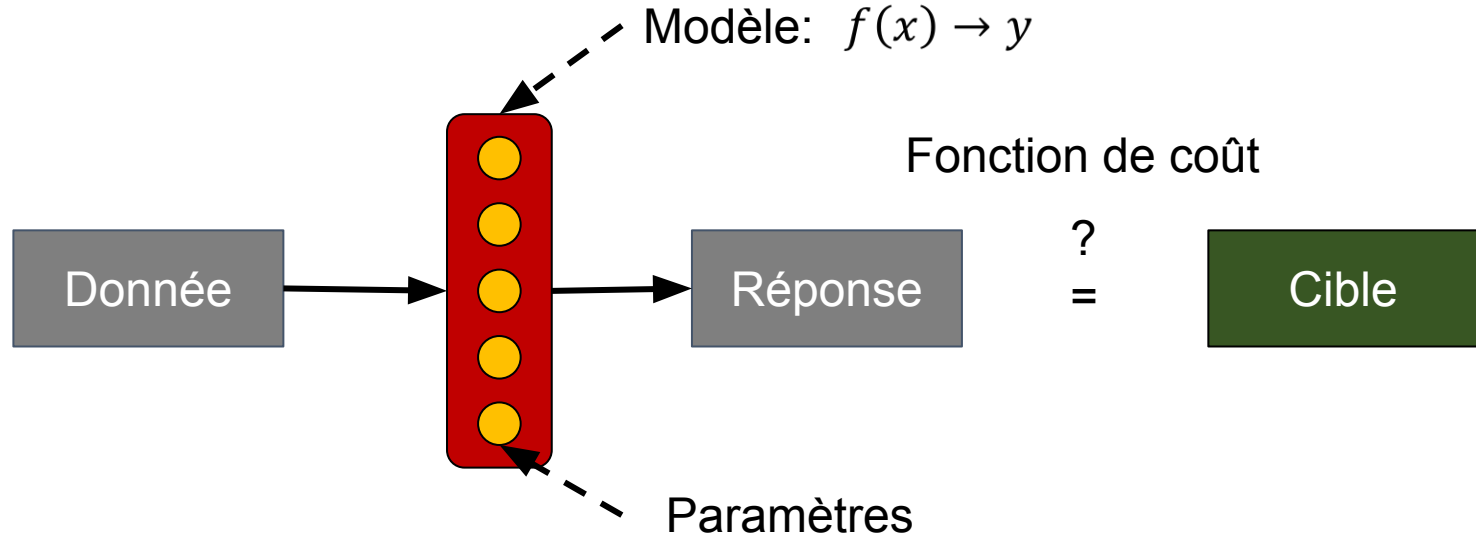


Nouvel exemple

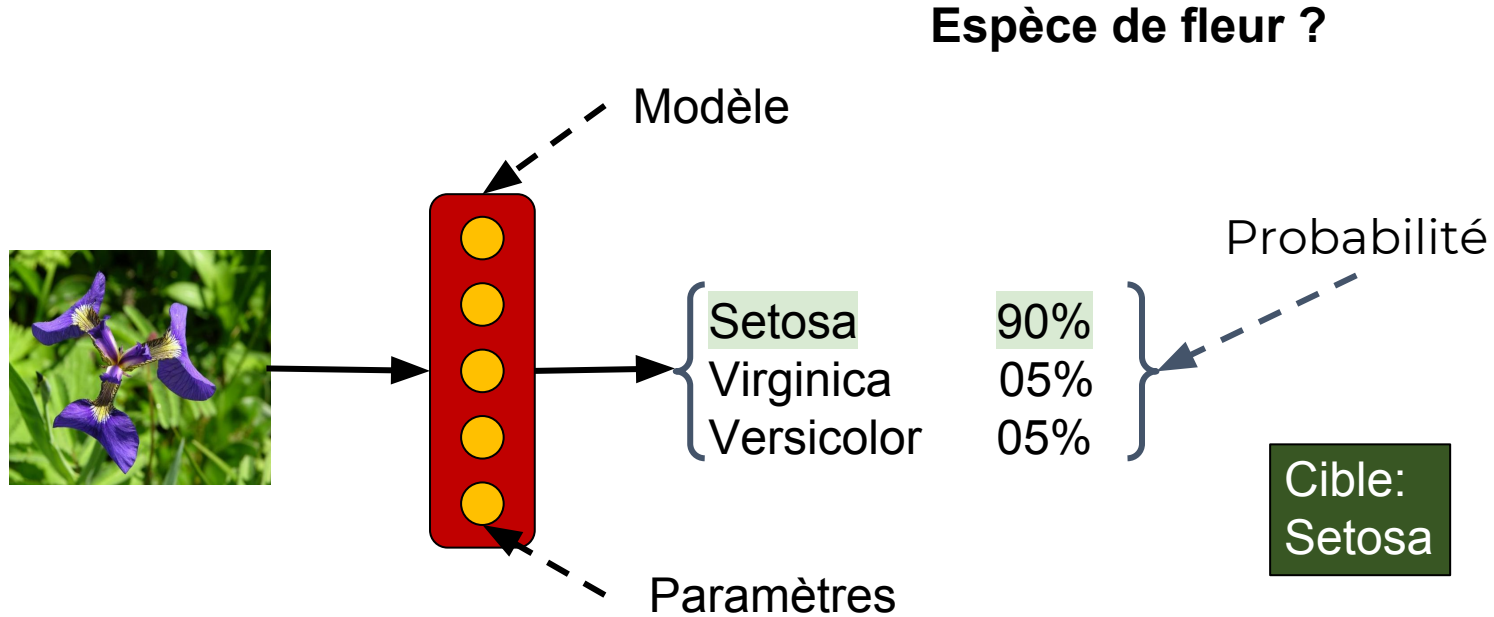


Types de modèles

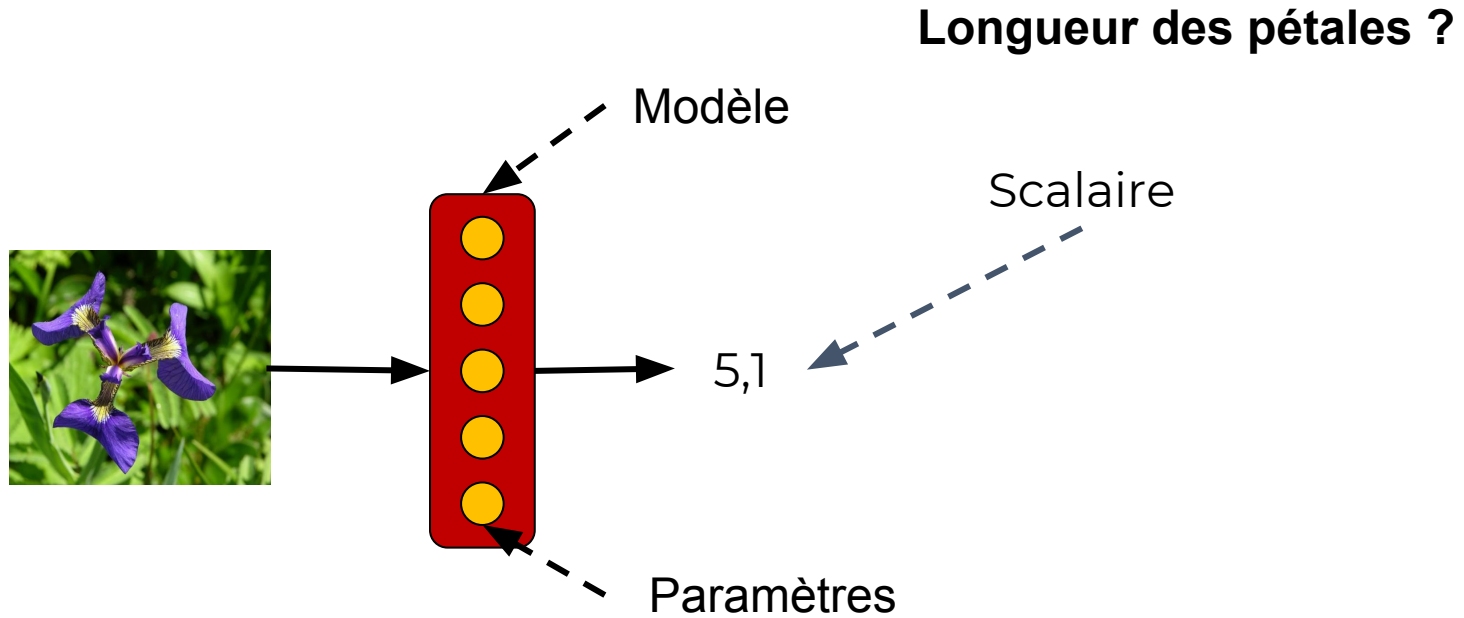
Définition d'un modèle



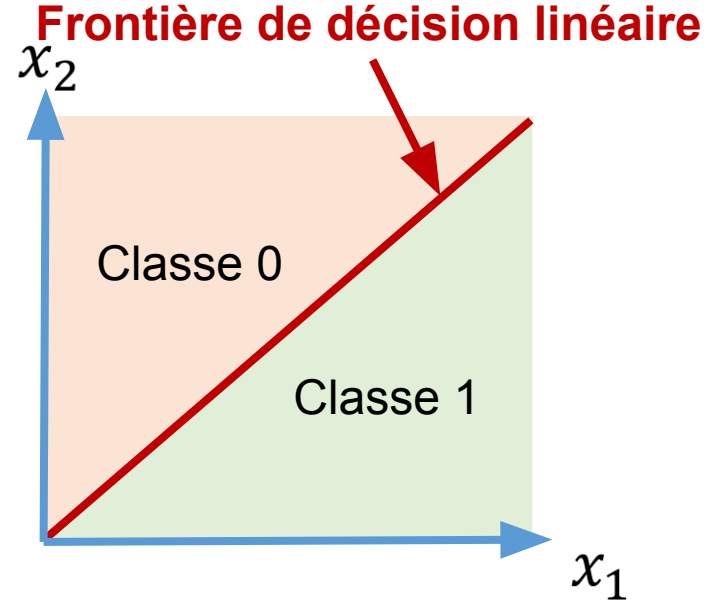
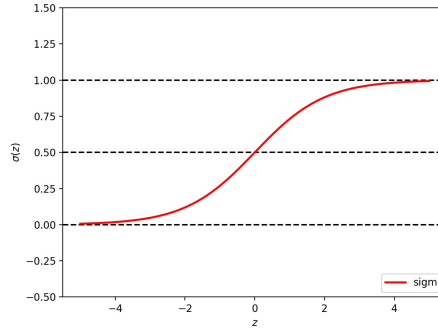
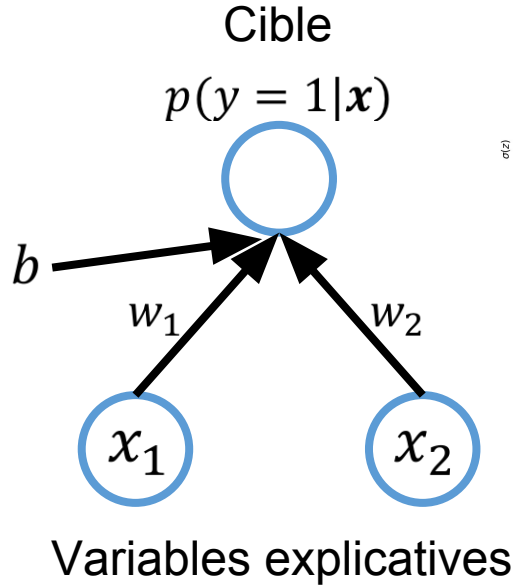
Cas de la classification



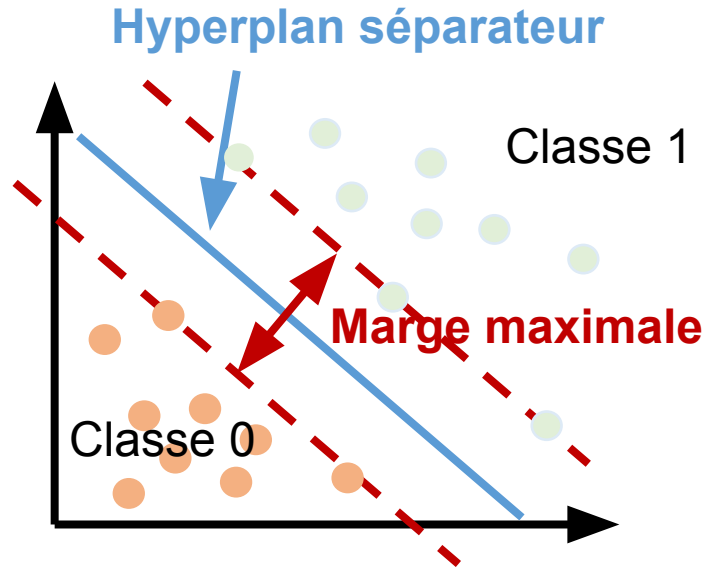
Cas de la régression



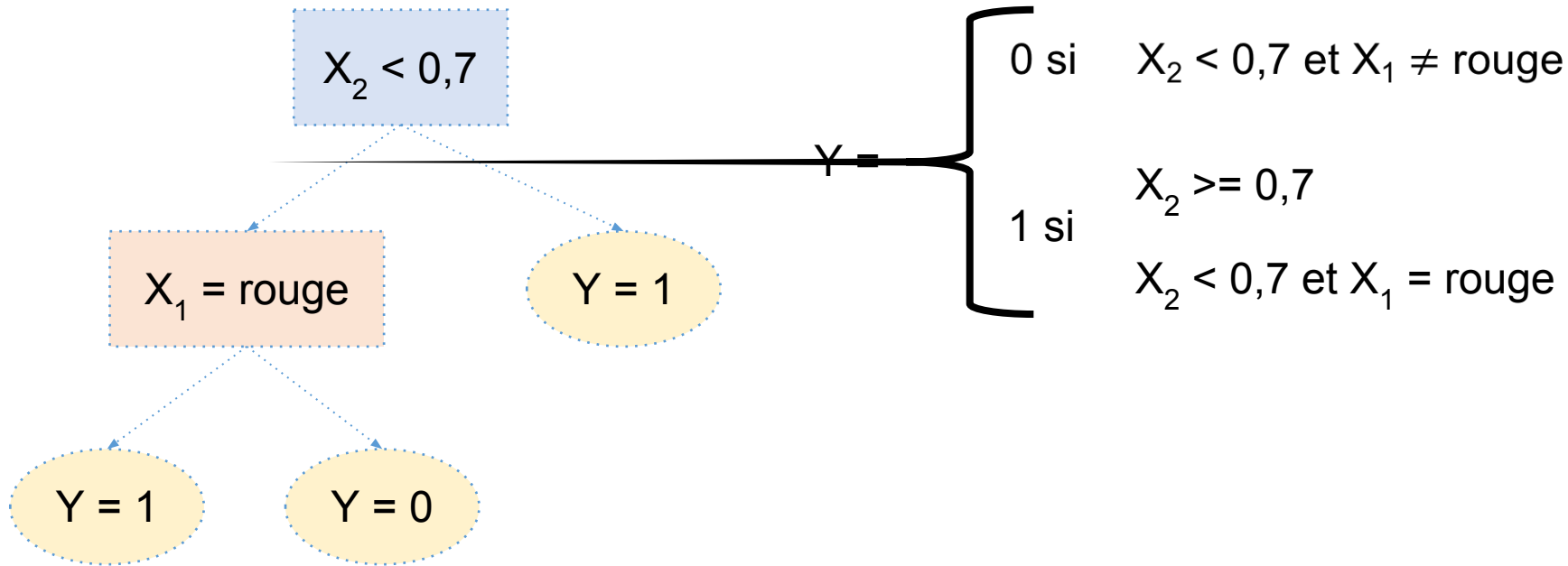
Modèles: Régression logistique



Machines à support de vecteur



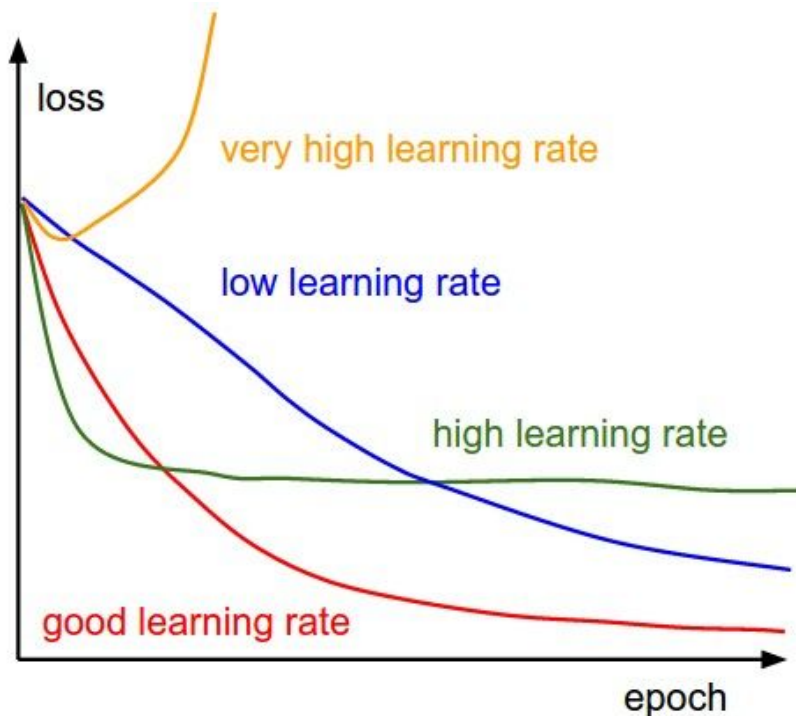
Modèles: Arbre de décision



Comment entraîner un modèle ?

Hyper-paramètres

- Paramètre dont la valeur est fixée avant l'apprentissage.
- Le bon choix d'hyper-paramètres a un impact fort sur la performance

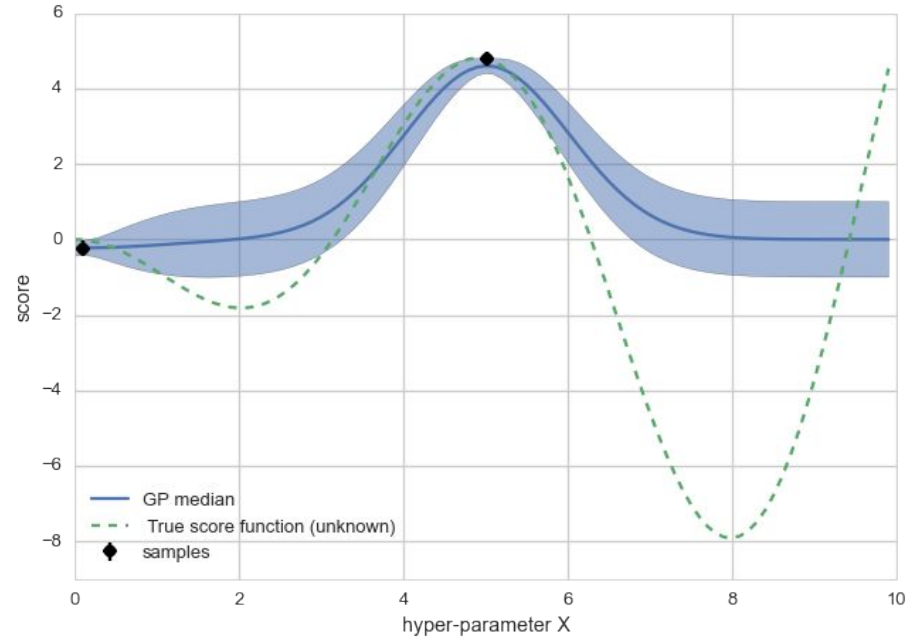


Hyper-paramètres: Exemples

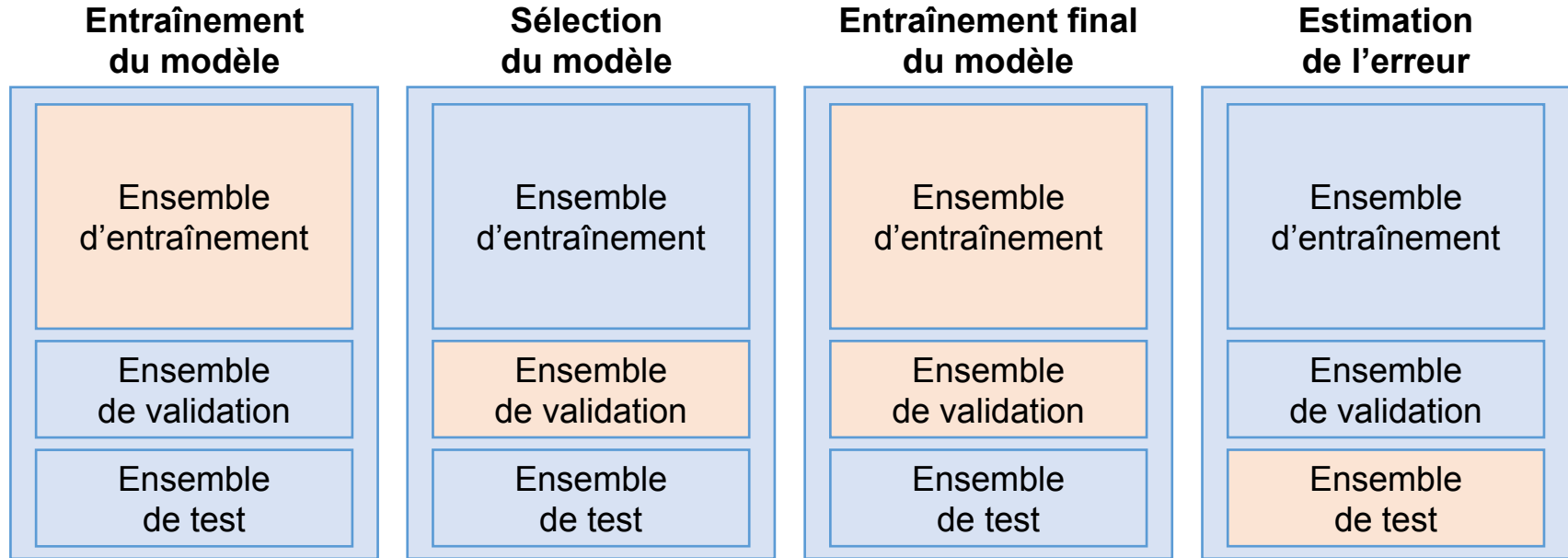
- Coefficient d'apprentissage
- Nombre de modèles (ensemble)
- Taille du modèle (réseaux)
- Poids des différents coûts.
- Taille des batches.
- Transformations d'entrée.
- Nombre maximal d'itérations.
- Choix d'algorithme d'apprentissage.
- Nombre de variables choisies en entrée.
- Type de modèle
- Type de couches (reseaux)
- Représentations choisies.

Choix des hyper-paramètres

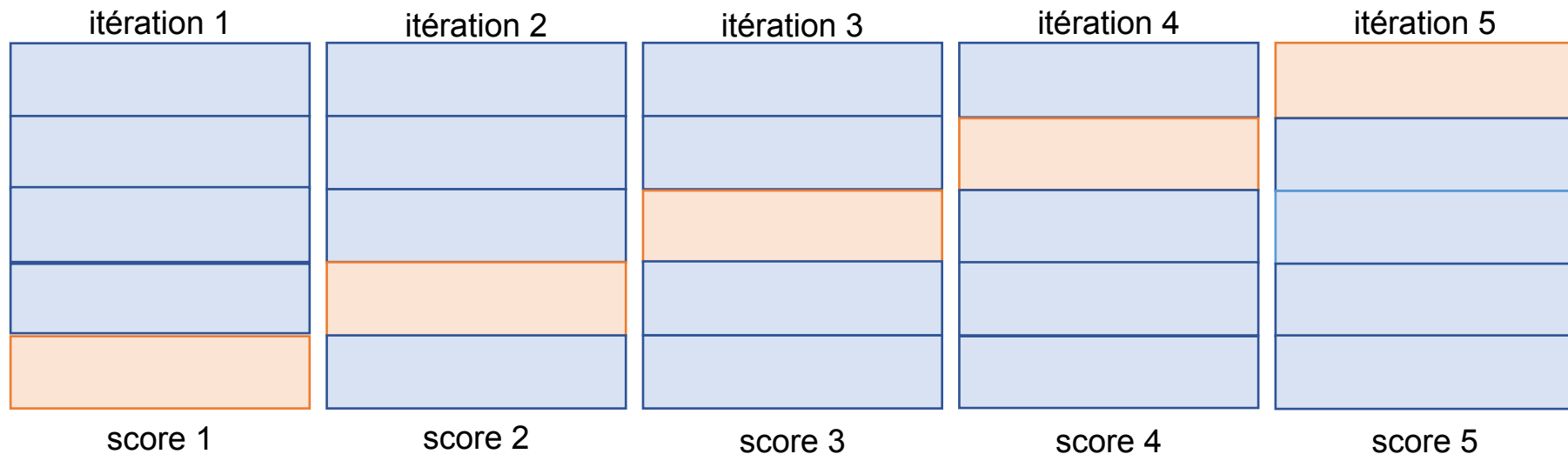
- Grid-search.
- Recherche aléatoire.
- Recherche par optimisation bayésienne.




Entraînement, validation et test




Validation croisée



Score final = moyenne(score 1, score 2, score 3, score 4, score 5)

 Ensemble d'entraînement et de validation

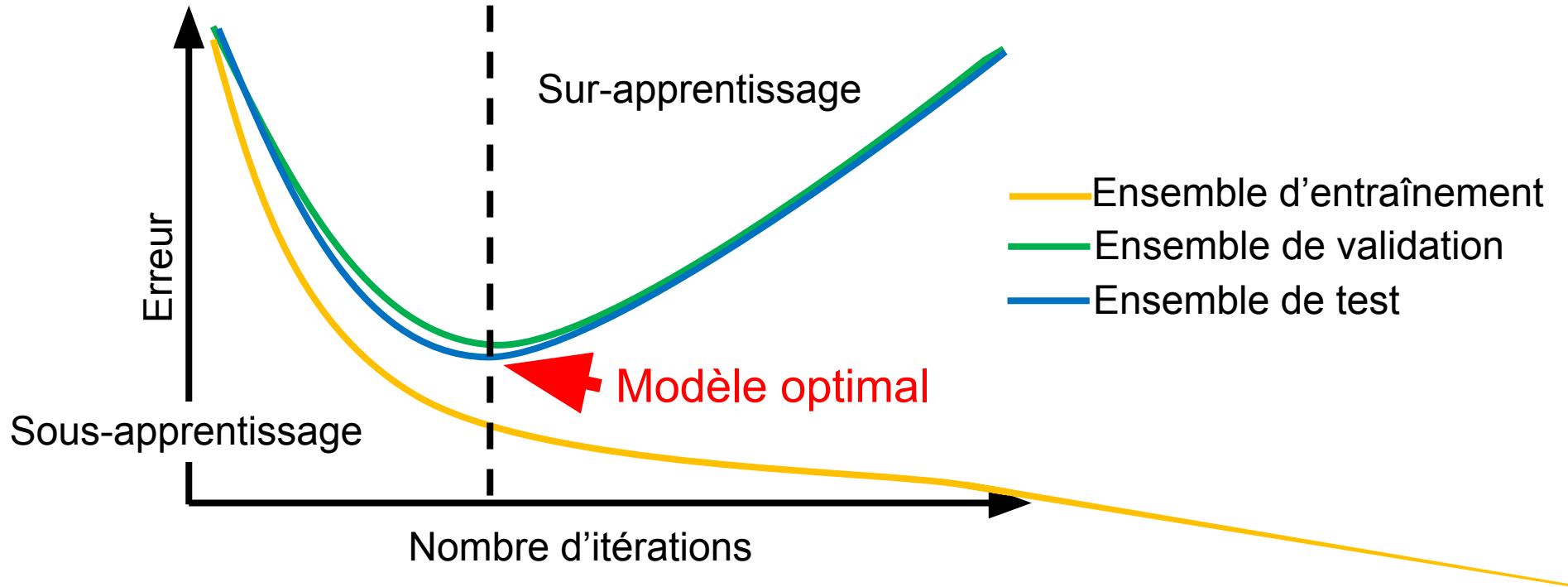
 Ensemble de test

Comment mieux généraliser?

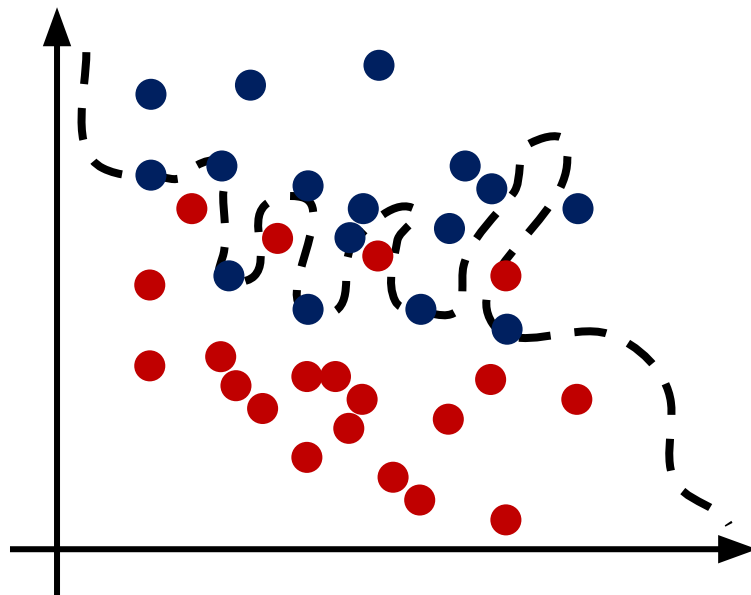
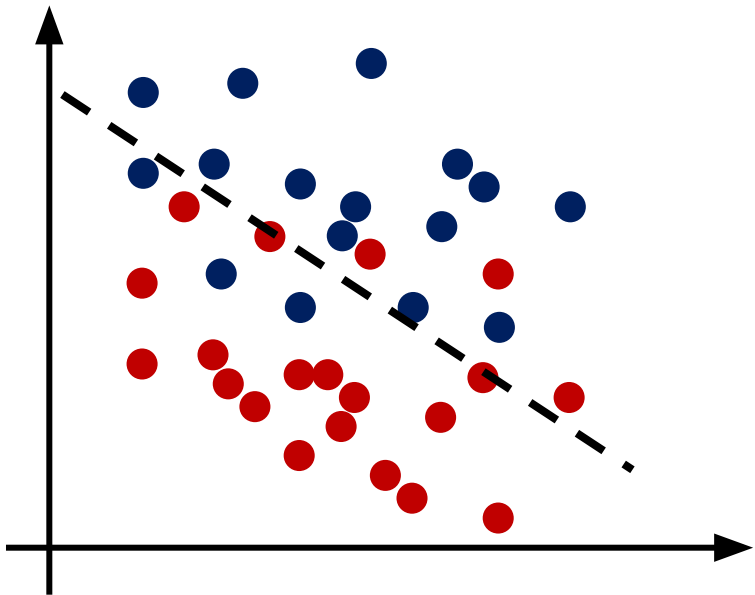
Généralisation

- La généralisation est la capacité du modèle a performer correctement sur de nouveaux exemples.
- Classiquement: la même tâche, mais aussi par exemple en Meta-Learning, une tâche différente.
- Essentiel: l'absence de généralisation est à la base des hivers de l'IA.

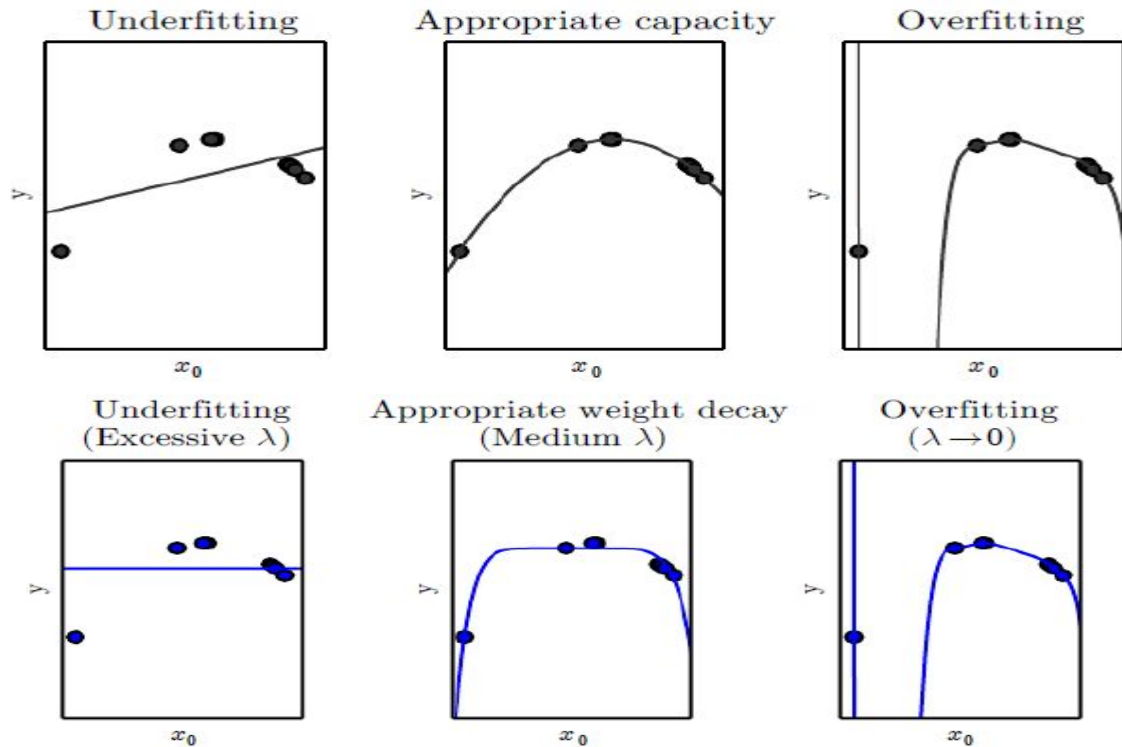
Modèles: Sur- et sous-apprentissage



Sur-apprentissage en classification

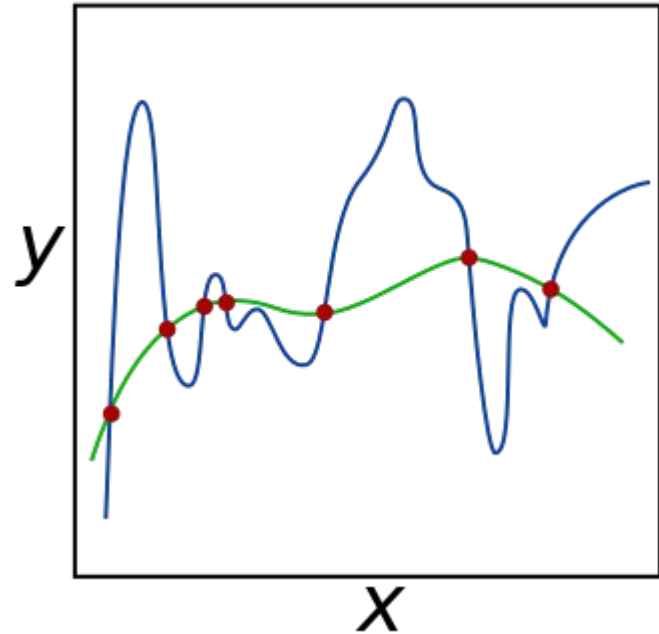


Régularisation (1)

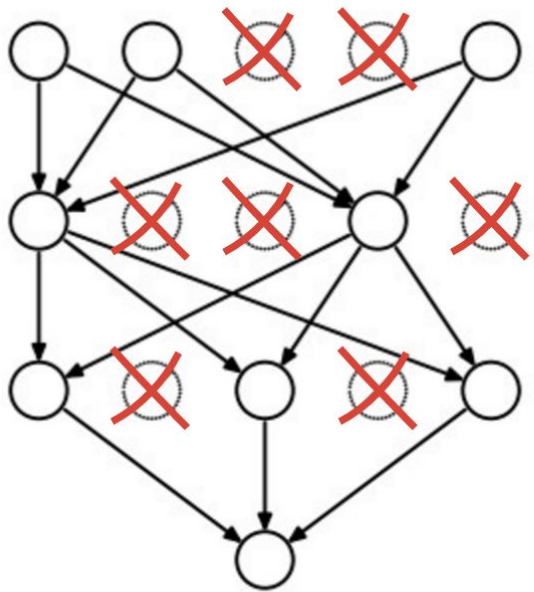


Régularisation (2)

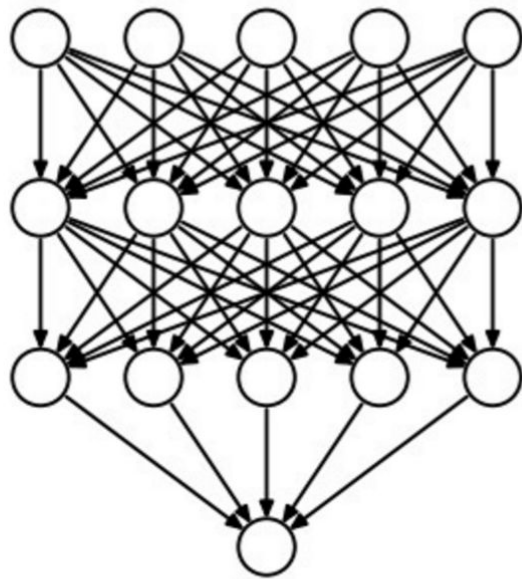
- La régularisation est un ensemble de techniques qui affectent l'apprentissage afin d'améliorer la performance sur le test.
- Contribue souvent à faire décroître la performance sur l'ensemble d'entraînement.
- C'est une méthode essentielle pour éviter le sur-apprentissage.



Examples (1): Dropout



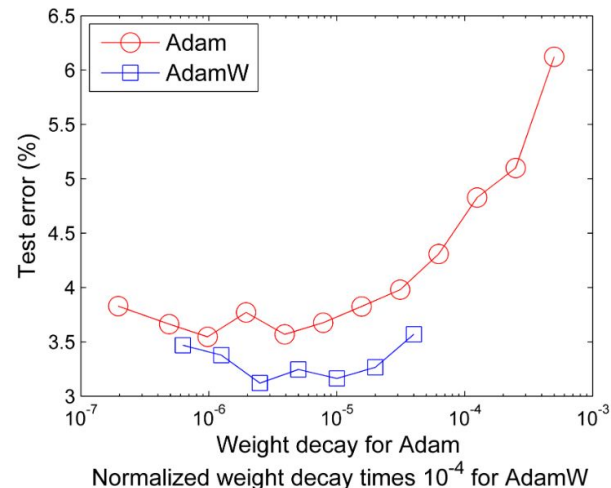
Training
 $p_{\text{keep}} = 0.75$



Test
 $p_{\text{keep}} = 1.0$

Exemples (2): Penalité

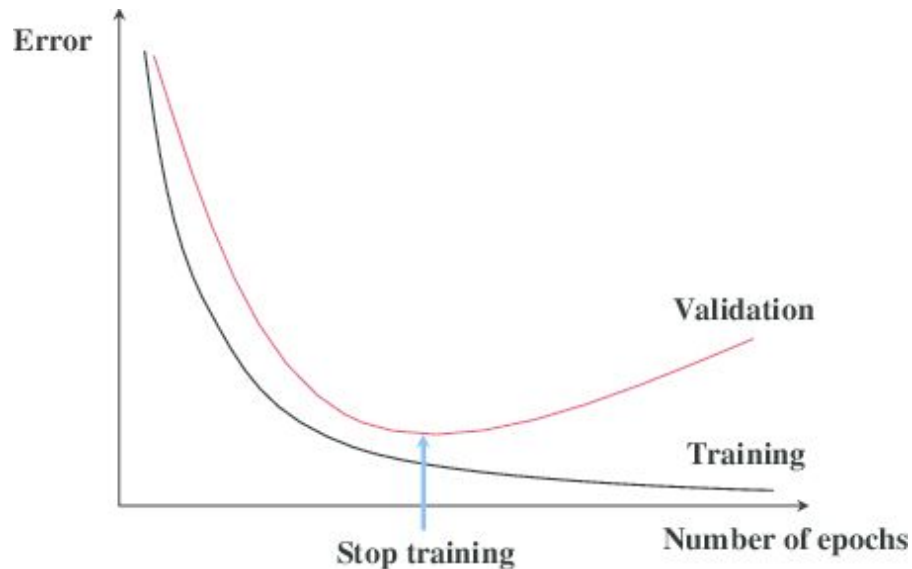
- Beaucoup de modèles cherchent à minimiser une fonction de coût.
- On peut ajouter à la fonction de coût un terme qui dépend du modèle.
- Un exemple est la pénalisation L2 sur les poids d'un modèle.



Effet d'une pénalité sur la précision

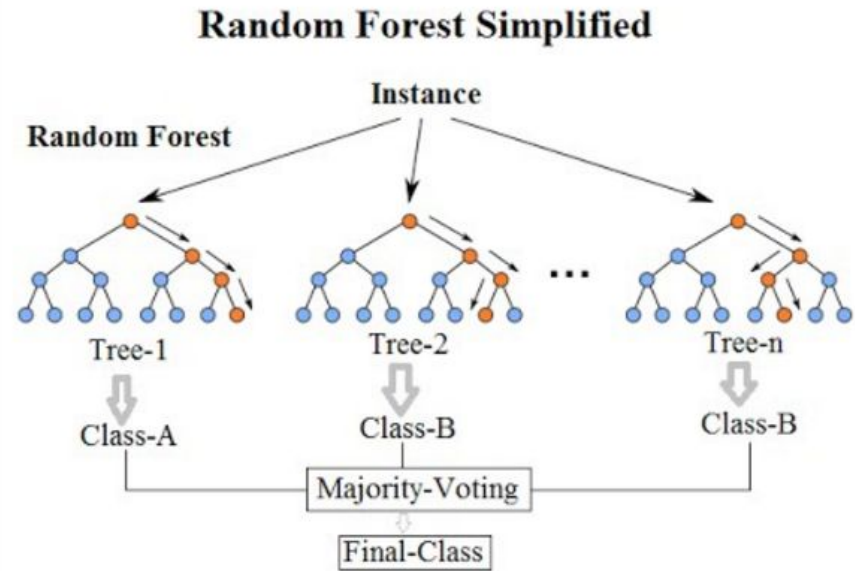
Exemples (3): Arrêt prématuré

- Pour des méthodes a gradient principalement, on arrête l'entraînement tôt
- En général, dès que l'erreur de validation arrête de baisser.



Méthodes d'ensemble

- Quitte à utiliser plus de ressources, on peut moyenner les prédictions de plusieurs modèles.
- Exemples: Random Forests, ensembles de réseaux neuronaux, ...



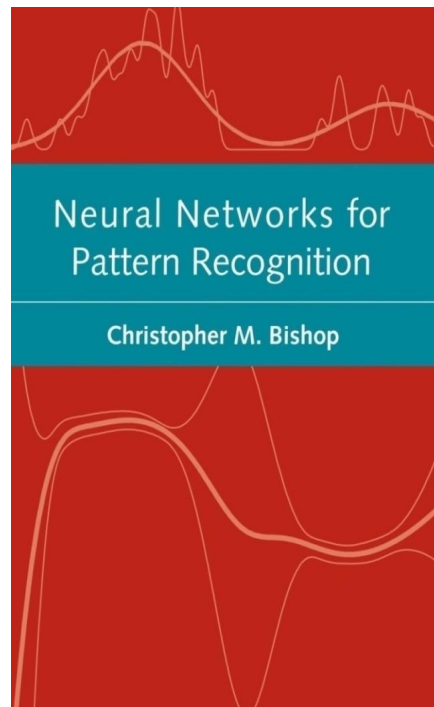
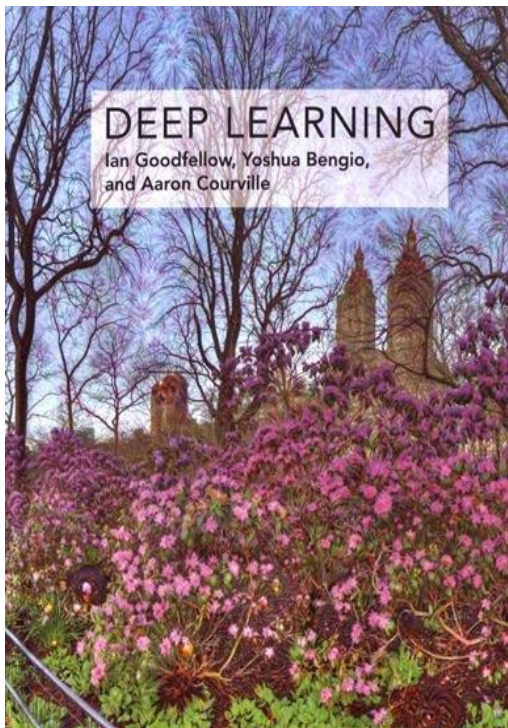
Conclusion

- Types de tâches et d'apprentissage.
- Préparation et nature des données.
- Types de modèles classiques.
- Choix des hyper-paramètres.
- Comment mieux généraliser.

À retenir

- Domaine très vaste, beaucoup de tâches et de types d'apprentissage.
- Il faut correctement préparer ses données et choisir les hyper-paramètres.
- Importance de la généralisation.

Pour aller plus loin



Contact

Tristan Sylvain

Doctorant au Mila, chercheur à Imagia

tristan.sylvain@gmail.com

<https://tisu32.github.io/about/>

