



Apprentissage automatique supervisé

Enseignante: Jamila Hamdi

Qu'est ce que l'apprentissage automatique (Machine Learning)?

Trois classes de problèmes d'apprentissage

Apprentissage Supervisé

Données : (x, y)

x est une donnée d'entrée, y est une étiquette (par exemple, une photo avec l'étiquette "chat")

Objectif: Apprendre à associer l'entrée à la sortie, c'est-à-dire $x \rightarrow y$

Un exemple: pour classifier Ceci est un chat



Un chat

Apprentissage Non Supervisé

Données: x ,

x est une donnée, il n'y a pas d'étiquettes !

Objectif: Apprendre la structure sous-jacente des données.

Un exemple: Comparaison, regroupement



Ces deux choses se ressemblent

Apprentissage par Renforcement

Données: pas de données, seulement des paires état-action (s, a)

Objectif: Maximiser la récompense future sur plusieurs étapes temporelles.

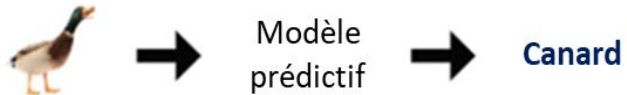
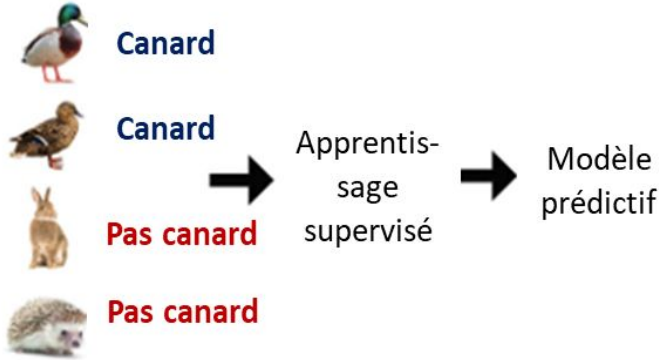
Un exemple: récompense = joie



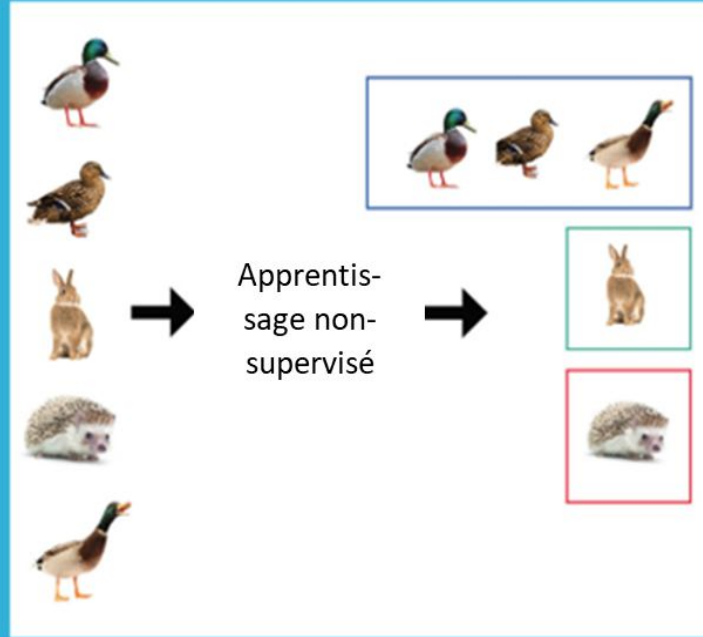
L'interaction avec le chat donne de la joie



Apprentissage Supervisé



Apprentissage non Supervisé



Qu'est ce que l'apprentissage automatique supervisé?

L'apprentissage est dit supervisé lorsque les données qui entrent dans le processus sont déjà **catégorisées** et que les algorithmes doivent s'en servir pour prédire un résultat en vue de pouvoir le faire plus tard lorsque les données ne seront plus catégorisées.



Classification/Régression

Fondamentalement, la classification consiste à prédire une étiquette et la régression consiste à prédire une quantité.

- **La classification** est le problème de prédiction d'une sortie d'étiquette de classe **discrète** pour un exemple.
- **La régression** est le problème de prédiction d'une **quantité continue** en sortie pour un exemple.



Exemples d'apprentissage supervisé



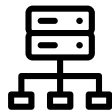
Analytique prédictive

Les systèmes analytiques prédictives permettent aux dirigeants d'entreprise à justifier leurs décisions ou à pivoter au profit de l'organisation.



Analyse des sentiments des clients

les entreprises peuvent extraire et classer des éléments d'information importants à partir de grands volumes de données, y compris le contexte, l'émotion et l'intention, avec très peu d'intervention humaine



Détection du spam

les organisations peuvent entraîner des bases de données à reconnaître des modèles ou des anomalies dans les nouvelles données afin d'organiser efficacement les courriers indésirables ou non.



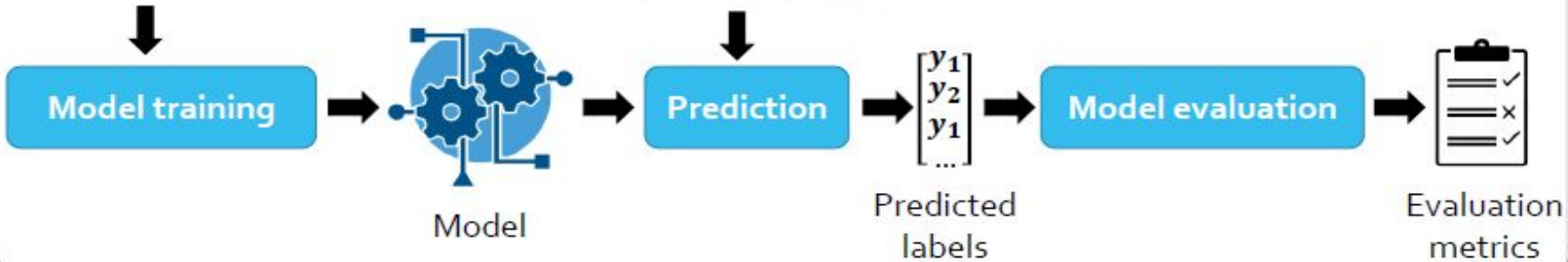
Reconnaissance d'images et d'objets

localiser, isoler et catégoriser des objets dans des vidéos ou des images, ce qui les rend utiles lorsqu'ils sont appliqués à diverses techniques de vision par ordinateur et d'analyse d'images.

[illegible]

Part V	# absent	CR	CR in time	CR quality	Test
10.15	3	6	Yes	Average	Pass
10.15	4	1	No	Bad	Fail
9.33	0	1	No	Bad	Fail
12.35	0	6	Yes	Good	Pass
16.78	1	6	No	Very Good	Pass
14.55	2	6	No	Very Good	Pass

Pctry Grade	# absent	CR	CR in time	CR quality	Test
30-35	3	6	Yes	Average	Pass
30-35	4	1	No	Bad	Fail
9-33	0	2	No	Bad	Fail
32-35	0	6	Yes	Good	Pass






Exemple d'apprentissage supervisé

© Davtechlab
www.davtechlab.com

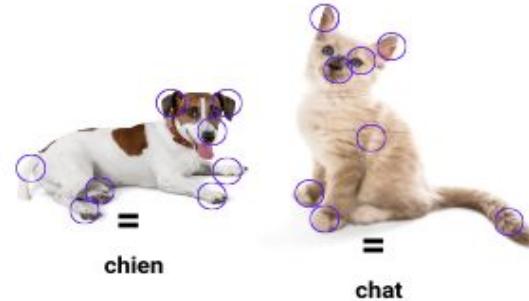
1

Phase d'apprentissage
du système

Les experts présentent au système des exemples
(données) étiquetés

exemples	étiquettes
	poisson
	chat
	chien

Le système détermine un **modèle** à partir des
données étiquetées



2

Phase de test du
système

nouvelle donnée

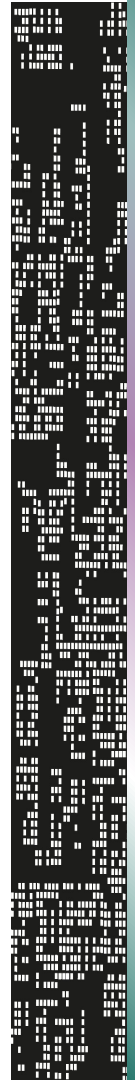


modèle préalablement appris

prédiction



résultat de la prédiction



Un aperçu des algorithmes d'apprentissage supervisé

Arbres de décision

Régression logistique

Régression linéaire



Random Forest

Machines à vecteurs de support
(SVM)

Naive Bayes

Apprentissage Supervisé : Régression



- la régression linéaire est utilisée pour identifier la relation entre une variable dépendante et une ou plusieurs variables indépendantes et est généralement utilisée pour faire des prédictions sur les résultats futurs.
- Lorsqu'il n'y a qu'une seule variable indépendante et une seule variable dépendante, on parle de régression linéaire simple. Lorsque le nombre de variables indépendantes augmente, on parle de régression linéaire multiple.
- Pour chaque type de régression linéaire, on cherche à tracer une ligne de meilleur ajustement, qui est calculée par la méthode des moindres carrés. Cependant, contrairement aux autres modèles de régression, cette ligne est droite lorsqu'elle est tracée sur un graphique.