



# **Module**

**MLCI**



**ML**

**CI**

**Module**

**MLCI**

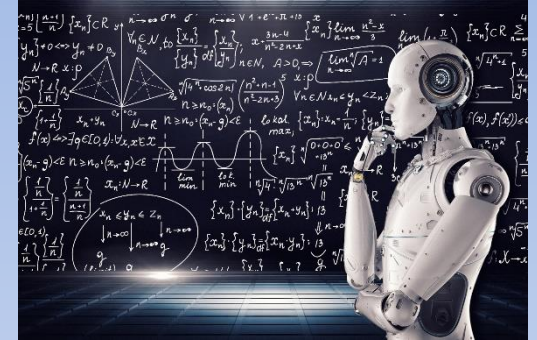
**ML**

**CI**

**Machine Learning**

**Computational  
Intelligence**

## **Module**



# **Machine Learning and Computational Intelligence**

**MLCI**

**Unité d'enseignement: UEF3**

**Crédit: 5**

**Coefficient: 3**

**Cours: 1H30/semaine**

**TP: 1H30/semaine**

**Dr. Fergani**

**Baha.fergani@univ-constantine2.dz**

# **Module**

**MLCI**

**Machine Learning**

**Deep Learning**

# **Module**

**MLCI**

**Machine Learning**

**Apprentissage par  
renforcement**

# **Module**

**MLCI**

**Machine Learning**

**Logique Floue**

# **Module**

**MLCI**

**Machine Learning**

**Computational intelligence**



# **Apprentissage automatique (Machine Learning)**

# **Apprentissage automatique (Machine Learning)**

**Apprentissage**

# **Apprentissage automatique (Machine Learning)**

**Apprentissage**

**Automatique**

# Apprentissage

L'apprentissage:

- Est un **comportement humain** naturel
- Qui est devenu un aspect essentiel des machines.

# Apprentissage automatique

- Apprentissage automatique est un champ d'étude de l'intelligence artificielle.
- Il vise à donner aux ordinateurs la capacité d'apprendre à partir des **données** sans **assistance**.

# Apprentissage automatique

- Apprentissage automatique est un champ d'étude de l'intelligence artificielle.
- Il vise à donner aux ordinateurs la capacité d'apprendre à partir des **données** sans **assistance**.

Données ???

# Apprentissage automatique

- L'apprentissage automatique nécessite deux ensembles de données:
  - **Ensemble de données pour l'entraînement:** c'est les données utilisées pour entraîner l'algorithme d'apprentissage.

Pendant cette phase, les paramètres du modèle peuvent être réglés (ajustés) en fonction des performances obtenues.

# Apprentissage automatique

- L'apprentissage automatique nécessite deux ensembles de données:
  - **Ensemble de données pour l'entraînement:** c'est les données utilisées pour entraîner l'algorithme d'apprentissage.

Pendant cette phase, les paramètres du modèle peuvent être réglés (ajustés) en fonction des performances obtenues.



# Apprentissage automatique

- L'apprentissage automatique nécessite deux ensembles de données:
  - **Ensemble de données pour le test:** il est utilisé pour évaluer les performances du modèle sur les données non-vues.

# Types d'apprentissage automatique

# Types d'apprentissage automatique

Les méthodes d'apprentissage peuvent être classées en **trois** principales catégories:

# Types d'apprentissage automatique

Les méthodes d'apprentissage peuvent être classées en **trois** principales catégories:

**Apprentissage  
supervisé**



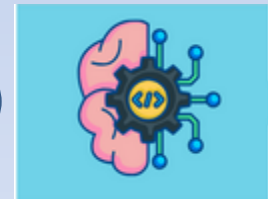
# Types d'apprentissage automatique

Les méthodes d'apprentissage peuvent être classées en **trois** principales catégories:

**Apprentissage  
supervisé**



**Apprentissage non-  
supervisé**



# Types d'apprentissage automatique

Les méthodes d'apprentissage peuvent être classées en **trois** principales catégories:

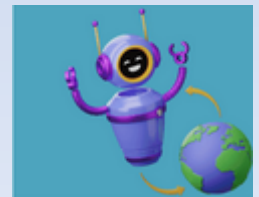
**Apprentissage  
supervisé**



**Apprentissage non-  
supervisé**



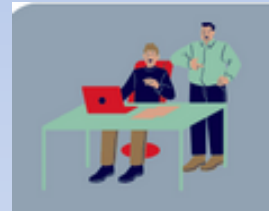
**Apprentissage par  
renforcement**



# Types d'apprentissage automatique

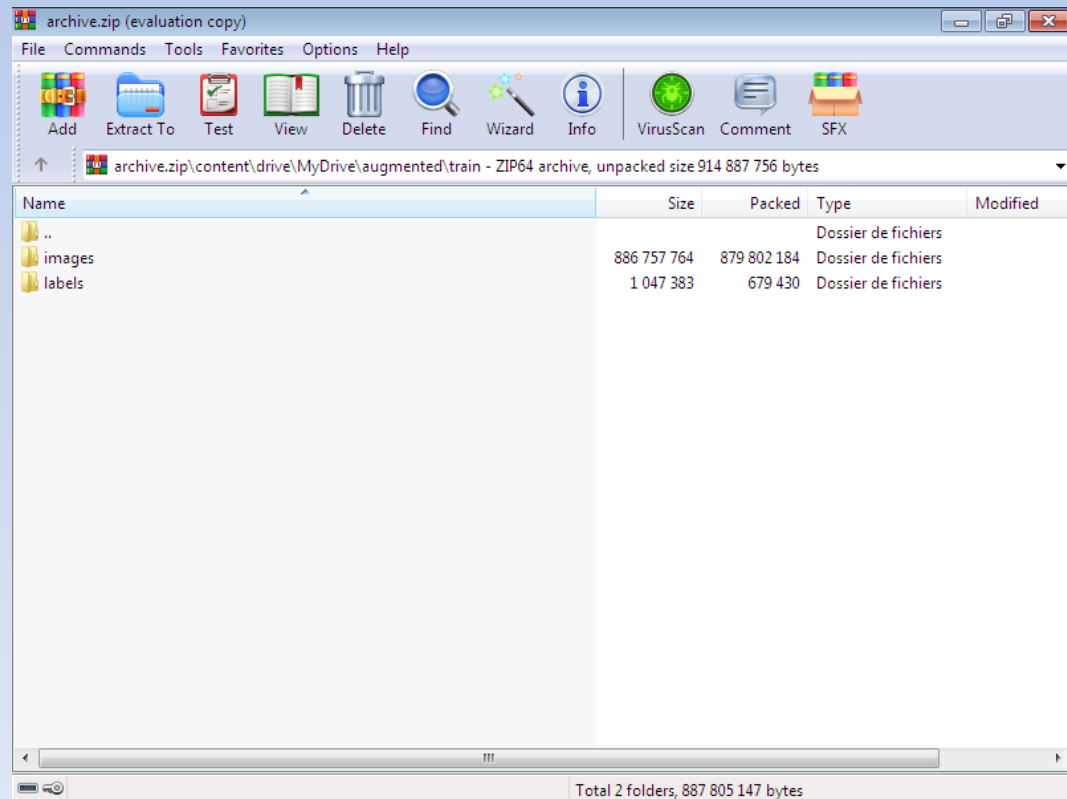
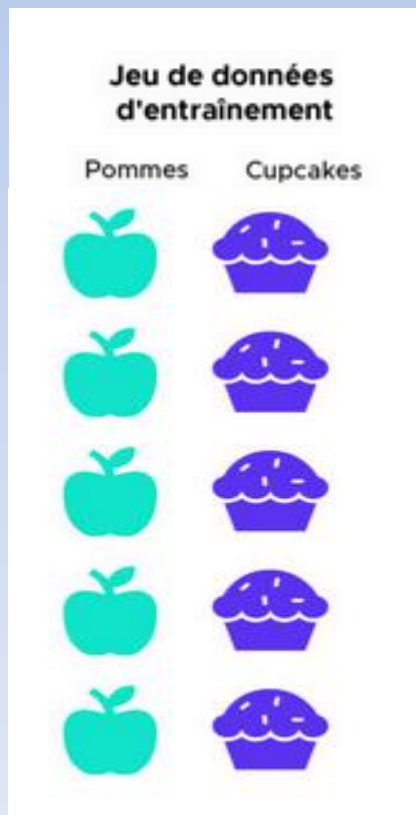
Les méthodes d'apprentissage peuvent être classées en **trois** principales catégories:

**Apprentissage  
supervisé**



# Apprentissage supervisé

- L'apprentissage supervisé (supervised learning) s'intéresse aux données étiquetées.





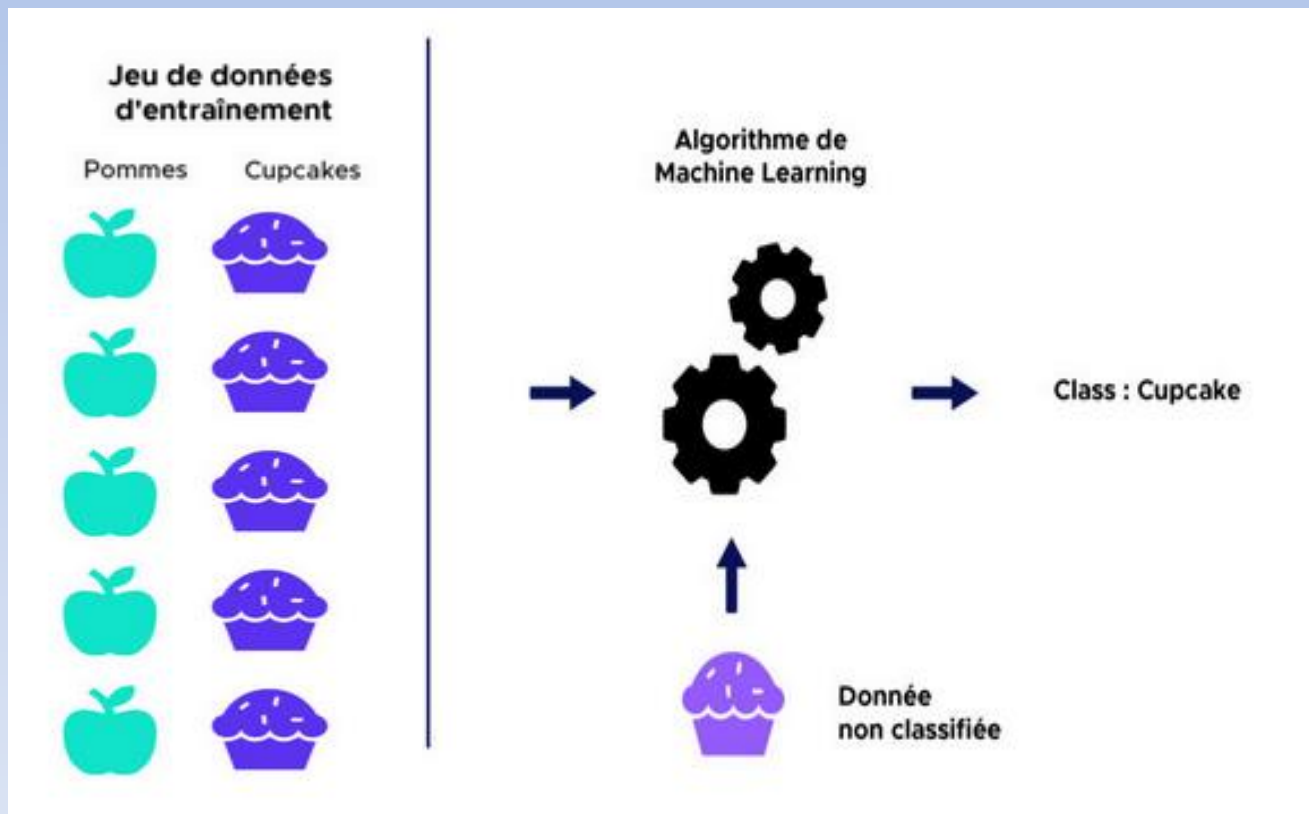
# Apprentissage supervisé

## Objectif:

- Est de prédire l'étiquette inconnu  $y$
- Associée à une nouvelle observation  $x$ ,
- A partir de la connaissance fournie par les  $N$  observations étiquetées du jeu de données.

# Apprentissage supervisé

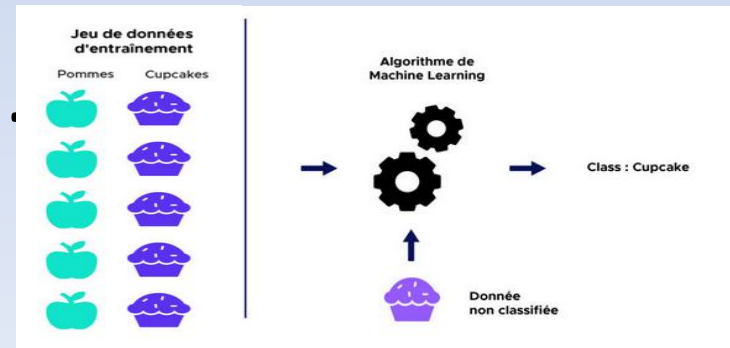
## Objectif:



# Apprentissage supervisé

Lors de l'apprentissage supervisé, l'algorithme reçoit:

- Un ensemble de données qui est étiqueté
- Sur lequel il va pouvoir s'entraîner **et** définir un modèle de prédiction.



# Apprentissage supervisé

- Cet algorithme pourra par la suite être utilisé sur de nouvelles données.
- Afin de prédire leurs valeurs de sorties correspondantes.

# Apprentissage supervisé

**Exemples d'algorithmes d'apprentissage**

**supervisé:** les algorithmes d'apprentissage

supervisé les plus utilisés sont :

- Support Vector Maching (SVM).
- L'arbre de décision.
- K-Nearest Neighbors (KNN).

# Apprentissage supervisé

**Exemples d'algorithmes d'apprentissage**

**supervisé:** les algorithmes d'apprentissage

supervisé les plus utilisés sont :

- Support Vector Maching (SVM).
- L'arbre de décision.
- K-Nearest Neighbors (KNN).

# Apprentissage supervisé

**Exemples d'algorithmes d'apprentissage**

**supervisé:** les algorithmes d'apprentissage

supervisé les plus utilisés sont :

- Support Vector Maching (SVM).
- L'arbre de décision.
- K-Nearest Neighbors (KNN).

# Apprentissage supervisé

**Exemples d'algorithmes d'apprentissage**

**supervisé:** les algorithmes d'apprentissage

supervisé les plus utilisés sont :

- Support Vector Maching (SVM).
- L'arbre de décision.
- K-Nearest Neighbors (KNN).



**Algorithme**  
***k* plus proches voisins**  
**(*k* Nearest Neighbors: KNN)**

# KNN

- L'algorithme KNN est utilisé pour la classification ou la régression.
- En classification, l'algorithme détermine à quelle classe appartient un échantillon en fonction de ses voisins les plus proches.

# KNN

- L'algorithme KNN est utilisé pour la classification ou la régression.
- En classification, l'algorithme détermine à quelle classe appartient un échantillon en fonction de ses voisins les plus proches.

# KNN

- L'algorithme KNN est utilisé pour la classification ou la régression.
- En classification, l'algorithme détermine à quelle classe appartient un échantillon en fonction de ses voisins les plus proches.
- En régression, l'algorithme calcule la moyenne des valeurs cibles des  $k$  plus proches voisins.

# KNN

## Principe

- L'algorithme kNN suppose que des objets similaires existent à proximité.

En d'autres termes, les éléments similaires sont proches les uns des autres.

# KNN

## Principe

- L'algorithme kNN suppose que des objets similaires existent à proximité.

En d'autres termes, les éléments similaires sont proches les uns des autres.

# Etapes de KNN

## Algorithme

# KNN

## Algorithme

**Etape 1:** Charger les données

**Etape 2:** Initialiser **k**: le nombre de voisins.

**Etape 3:** Calculer toutes les distances entre cette observation en entrée et les autres observations du jeu de données,

**Etape 4:** Conserver les **k** observations du jeu de données qui sont les plus « proches » de l'observation à prédire,



# KNN

## Algorithme

**Etape 1:** Charger les données

**Etape 2:** Initialiser **k**: le nombre de voisins.

**Etape 3:** Calculer toutes les **distances** entre l'**observation en entrée** et les autres **observations** du jeu de données.

**Etape 4:** Conserver les **k** observations du jeu de données qui sont les plus « proches » de l'observation à prédire,

# KNN

## Algorithme

**Etape 1:** Charger les données

**Etape 2:** Initialiser  $k$ .

**Etape 3:** Calculer toutes les distances entre cette observation en entrée et les autres observations du jeu de données,

**Etape 4:** Conserver les  $k$  observations du jeu de données qui sont les plus « **proches** » de l'observation à prédire,

# KNN

## Algorithme

**Etape 5:** Prendre les valeurs des observations retenues:

- Si on effectue une **régression**: l'algorithme calcule la moyenne (ou la médiane) des valeurs des observations retenues.
- Si on effectue une **classification**, l'algorithme assigne une étiquette (label) de la classe majoritaire à la donnée qui était inconnue.

# KNN

## Algorithme

**Etape 5:** Prendre les valeurs des observations retenues:

- Si on effectue une **régression**: l'algorithme calcule la moyenne (ou la médiane) des valeurs des observations retenues.
- Si on effectue une **classification**, l'algorithme assigne une étiquette (label) de la classe majoritaire à la donnée qui était inconnue.

# KNN

## Algorithme

**Etape 5:** Prendre les valeurs des observations retenues:

- Si on effectue une **régression**: l'algorithme calcule la moyenne (ou la médiane) des valeurs des observations retenues.
- Si on effectue une **classification**, l'algorithme assigne une étiquette (label) de la classe majoritaire à la donnée qui était inconnue.

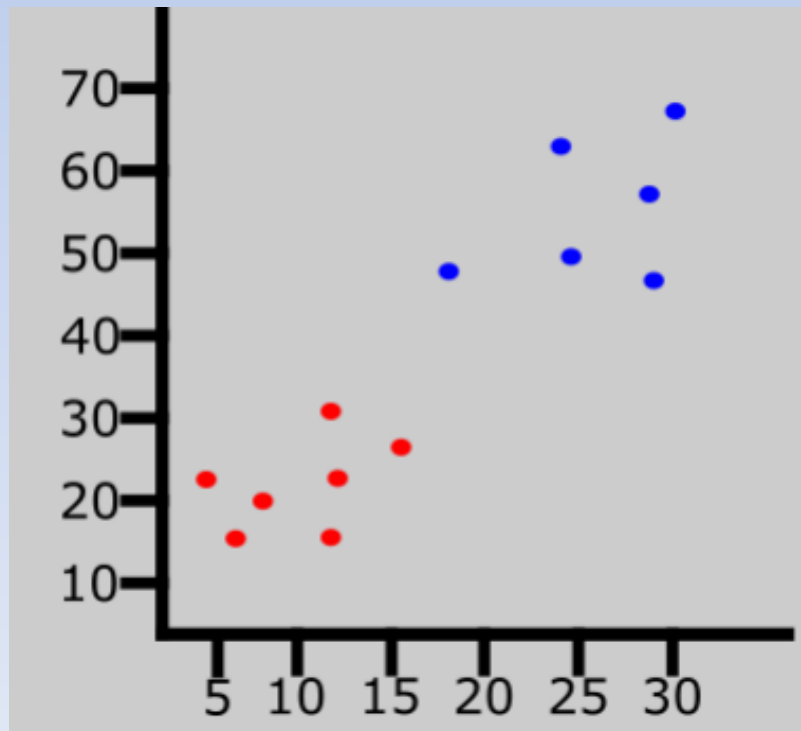
# KNN

## Algorithme

**Etape 1:** Charger les données.

Un ensemble de données regroupées en deux groupes: rouge et bleu.

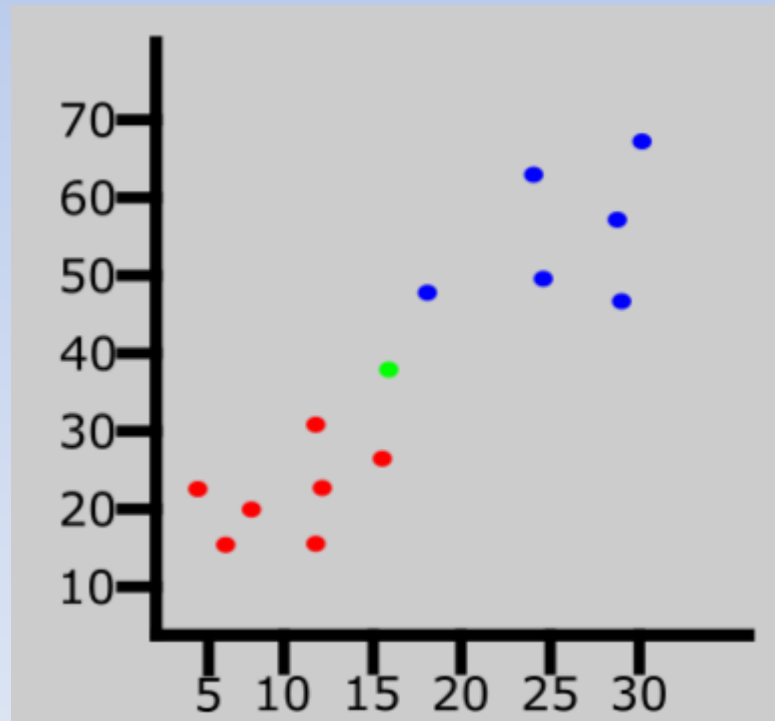
**Etape 2:**  $k=3$ .



# KNN

## Algorithme

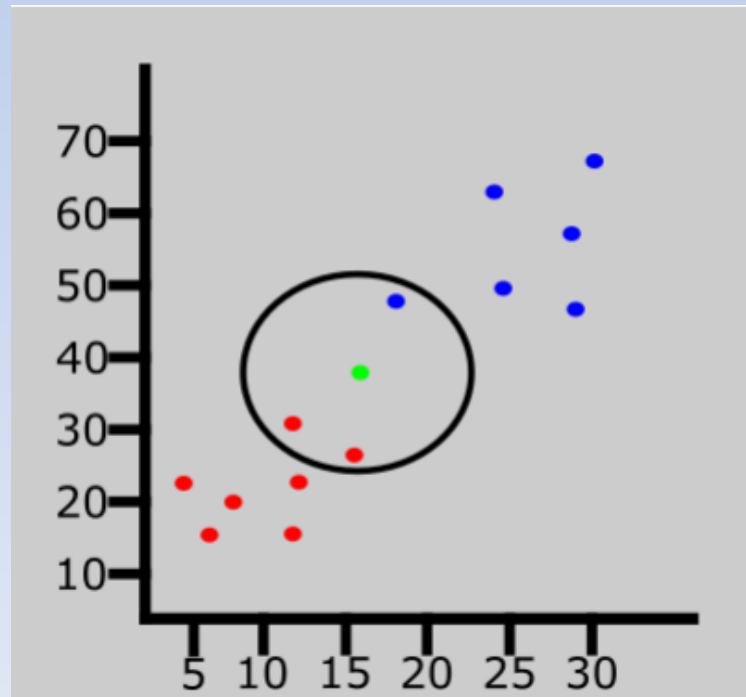
**But:** assigner le point **vert** à un des deux groupes.



# KNN

## Algorithme

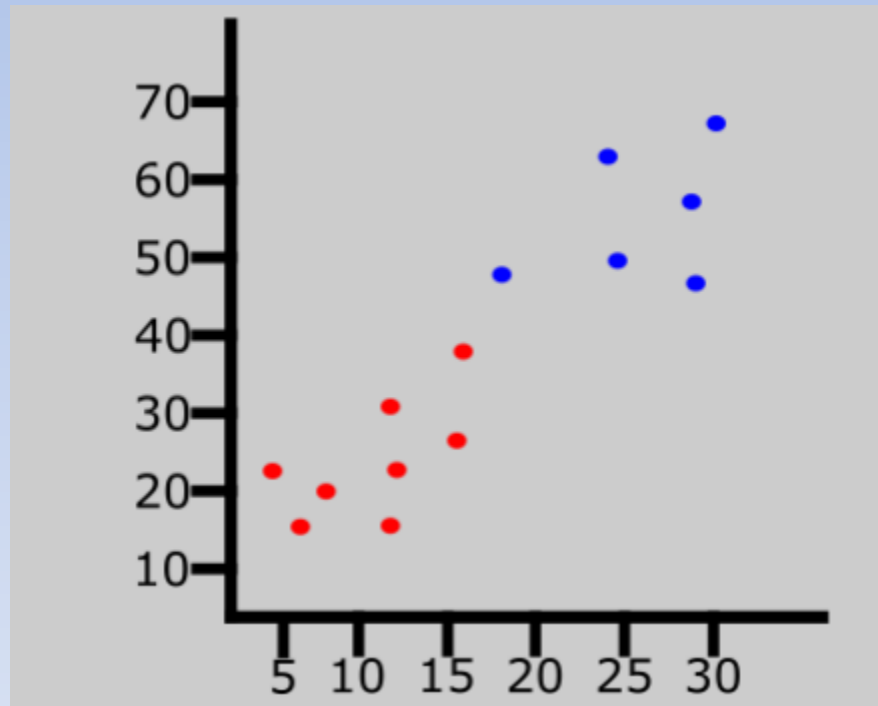
**Etape 4:** Conserver les **k=3** observations du jeu de données qui sont les plus « proches » de l'observation représentée par le point vert.





# KNN

## Algorithmme



# KNN

## Algorithme

## Remarques

# KNN

## Algorithme

Pour cet algorithme,

- Le choix du nombre **k**.
- et
- Le choix de la **fonction de similarité**

Sont des étapes qui peuvent conduire à une forte variabilité des résultats.

# KNN

## Algorithme

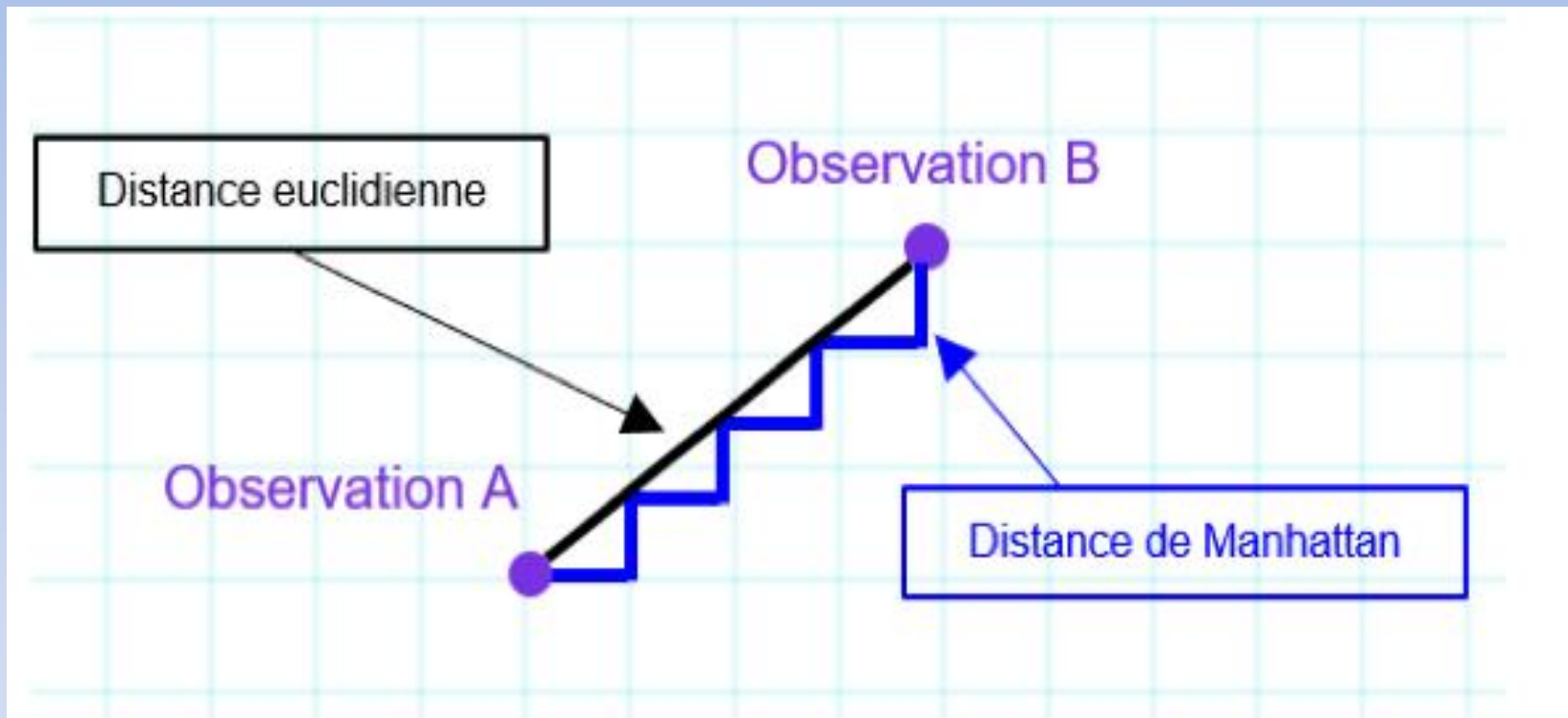
Pour cet algorithme,

- Le choix du nombre **k**.
- et
- Le choix de la **fonction de similarité**

Sont des étapes qui peuvent conduire à une forte variabilité des résultats.

# KNN

## Distance



$$\sum_{i=1}^n |x_i - y_i|$$

*Manhattan*

$$\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

*Euclidienne*

# KNN

## Avantages

# KNN

## Avantages

- L'algorithme est simple et facile à mettre en œuvre.
- Il n'est pas nécessaire de construire un modèle, d'ajuster plusieurs paramètres ou de faire des hypothèses supplémentaires.
- L'algorithme est polyvalent. Il peut être utilisé pour la classification, la régression et la recherche d'informations.

# KNN

## Avantages

- L'algorithme est simple et facile à mettre en œuvre.
- Il n'est pas nécessaire de construire un modèle, d'ajuster plusieurs paramètres ou de faire des hypothèses supplémentaires.
- L'algorithme est polyvalent. Il peut être utilisé pour la classification, la régression et la recherche d'informations.



# KNN

## Avantages

- L'algorithme est simple et facile à mettre en œuvre.
- Il n'est pas nécessaire de construire un modèle, d'ajuster plusieurs paramètres ou de faire des hypothèses supplémentaires.
- L'algorithme est **polyvalent**. Il peut être utilisé pour la **classification**, la **régression** et la recherche d'informations.

# KNN

## Inconvénients

# KNN

## Inconvénients

L'algorithme ralentit si:

- Le nombre d'observations et/ou de variables augmente.
- Parce que, **KNN** parcourt l'ensemble des observations pour calculer chaque distance.

# KNN

## Inconvénients

L'algorithme ralentit si:

- Le nombre d'observations et/ou de variables augmente.
- Parce que, **KNN** parcourt l'ensemble des observations pour calculer chaque distance.

# KNN

## Inconvénients

L'algorithme ralentit si:

- Le nombre d'observations et/ou de variables augmente.
- Parce que, **KNN** parcourt l'ensemble des observations pour calculer chaque distance.

# Types d'apprentissage automatique

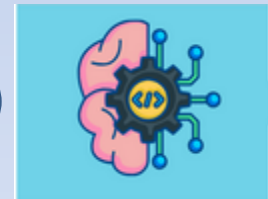
# Types d'apprentissage automatique

Les méthodes d'apprentissage peuvent être classées en trois principales catégories:

**Apprentissage  
supervisé**



**Apprentissage non-  
supervisé**



# Apprentissage non supervisé

- Aucun expert n'est disponible.
- L'algorithme doit découvrir par lui-même la structure des données.
- Le clustering est un algorithme d'apprentissage non supervisés.



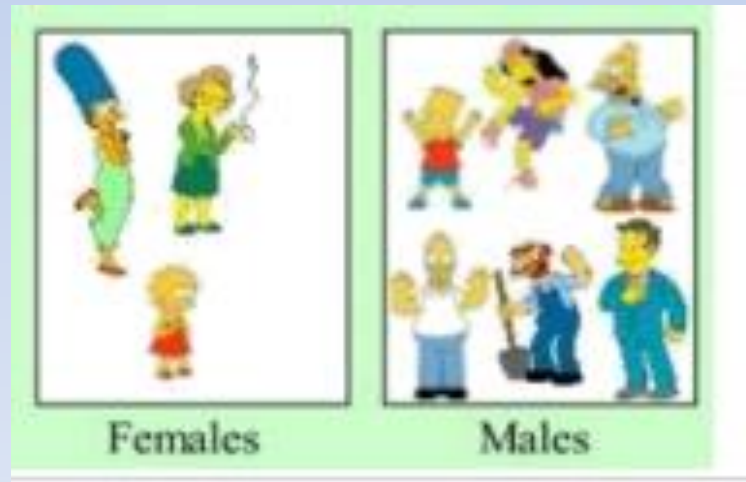
# Apprentissage non supervisé

- Aucun expert n'est disponible.
- L'algorithme doit découvrir par lui-même la structure des données.
- Le clustering est un exemple d'application des algorithmes d'apprentissage non supervisés.

# Apprentissage non supervisé

- Aucun expert n'est disponible.
- L'algorithme doit découvrir par lui-même la structure des données.
- Le clustering est un exemple d'application des algorithmes d'apprentissage non supervisés.

# Exemple de Clustering



# Exemple de Clustering

- Exemple de clustering



# Apprentissage supervisé / non-supervisé

	Apprentissage supervisé	Apprentissage non-supervisé
<b>Données d'entrée</b>	Données connues en entrée	Données inconnues en entrée
Complexité informatique	Complexe	Moins complexe
Domaines d'activités	Classification et régression	Clustering
Précision	Produit des résultats précis	Génère des résultats modérés

# Apprentissage supervisé / non-supervisé

	Apprentissage supervisé	Apprentissage non-supervisé
<b>Données d'entrée</b>	Données connues en entrée	Données inconnues en entrée
<b>Complexité informatique</b>	Complexe	Moins complexe
<b>Domaines d'activités</b>	Classification et régression	Clustering
<b>Précision</b>	Produit des résultats précis	Génère des résultats modérés

# Apprentissage supervisé / non-supervisé

	Apprentissage supervisé	Apprentissage non-supervisé
<b>Données d'entrée</b>	Données connues en entrée	Données inconnues en entrée
<b>Complexité informatique</b>	Complexe	Moins complexe
<b>Domaines d'activités</b>	Classification et régression	Clustering
<b>Précision</b>	Produit des résultats précis	Génère des résultats modérés

# Apprentissage supervisé / non-supervisé

	Apprentissage supervisé	Apprentissage non-supervisé
<b>Données d'entrée</b>	Données connues en entrée	Données inconnues en entrée
<b>Complexité informatique</b>	Complexe	Moins complexe
<b>Domaines d'activités</b>	Classification et régression	Clustering
<b>Précision</b>	Produit des résultats précis	Génère des résultats modérés



# Examples

# Exemples

## *Exemple 1:*

Supposons que l'on dispose d'une collection d'articles de journaux.

Comment identifier des groupes d'articles portant sur un même sujet?

# Exemples

- ***Exemple 1:*** discussion

On cherche à regrouper les articles portant sur un même sujet, sans disposer d'exemples d'articles dont on sait a priori qu'ils portent sur ce sujet, et sans connaître à l'avance les sujets à identifier.

On parlera donc de problème d'apprentissage non-supervisé.

# Exemples

## *Exemple 2:*

- Supposons que l'on dispose d'un certain nombre d'images représentant des chiens, et d'autres représentant des chats.
- Comment classer automatiquement une nouvelle image dans une des catégories « chien » ou « chat » ?

# Exemples

- ***Exemple 3:***

Supposons que l'on dispose d'une base de données regroupant les caractéristiques de logements dans une ville :

superficie, quartier, étage, prix, année de construction, nombre d'occupants, montant des frais de chauffage.

# Exemples

- *Exemple 3:*

Comment prédire la facture de chauffage à partir des autres caractéristiques pour un logement qui n'appartiendrait pas à cette base ?

# Exemples

- **Exemple 2 et exemple 3:**

Dans les exemples 2 et 3, on cherche à prédire une caractéristique qui est soit une catégorie (exemple 2), soit un montant de facture (exemple 3), à partir d'exemples pour lesquels on connaît la valeur de cette caractéristique. Il s'agit de problèmes d'apprentissage supervisé.

Algorithme de k-moyennes

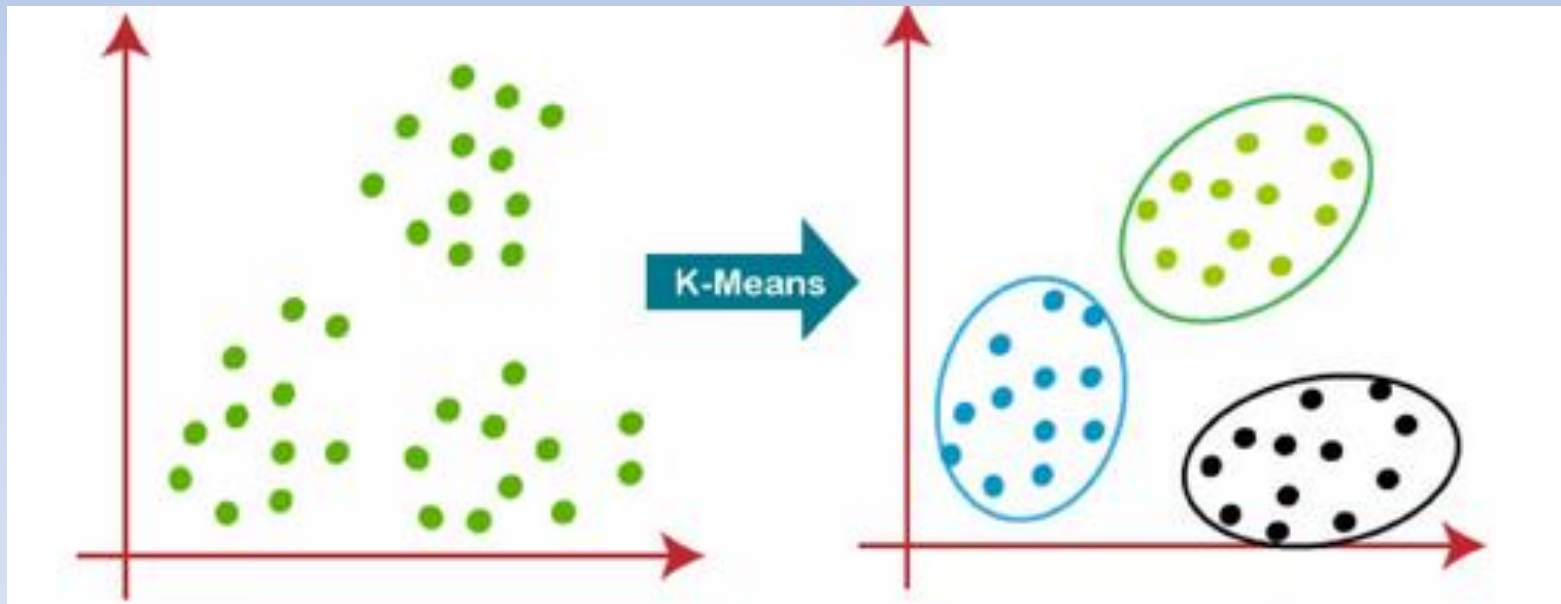
Algorithme des centres  
mobiles

K-means



# K-means

- **But:** assigner les éléments aux groupes



Avant k-means

Après K-means

Etapes

K-means

# K-means

## **Etape 1: initialisation des centroïdes**

Choisir aléatoirement **K** points.

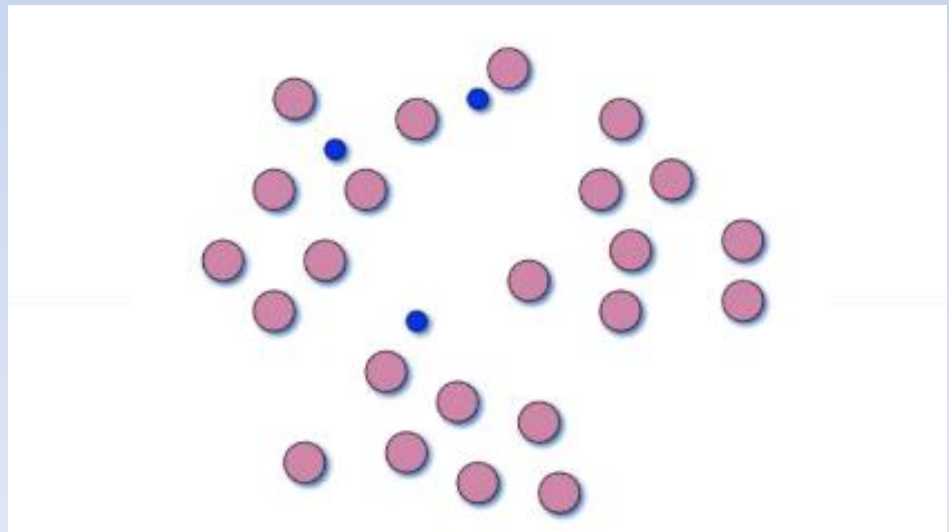
Ces points sont les centres des clusters initiaux (nommé centroïde).

# K-means

## Etape 1: initialisation des centroïdes

Choisir aléatoirement **K** points.

Ces points sont les centres des clusters initiaux (nommé centroïde).



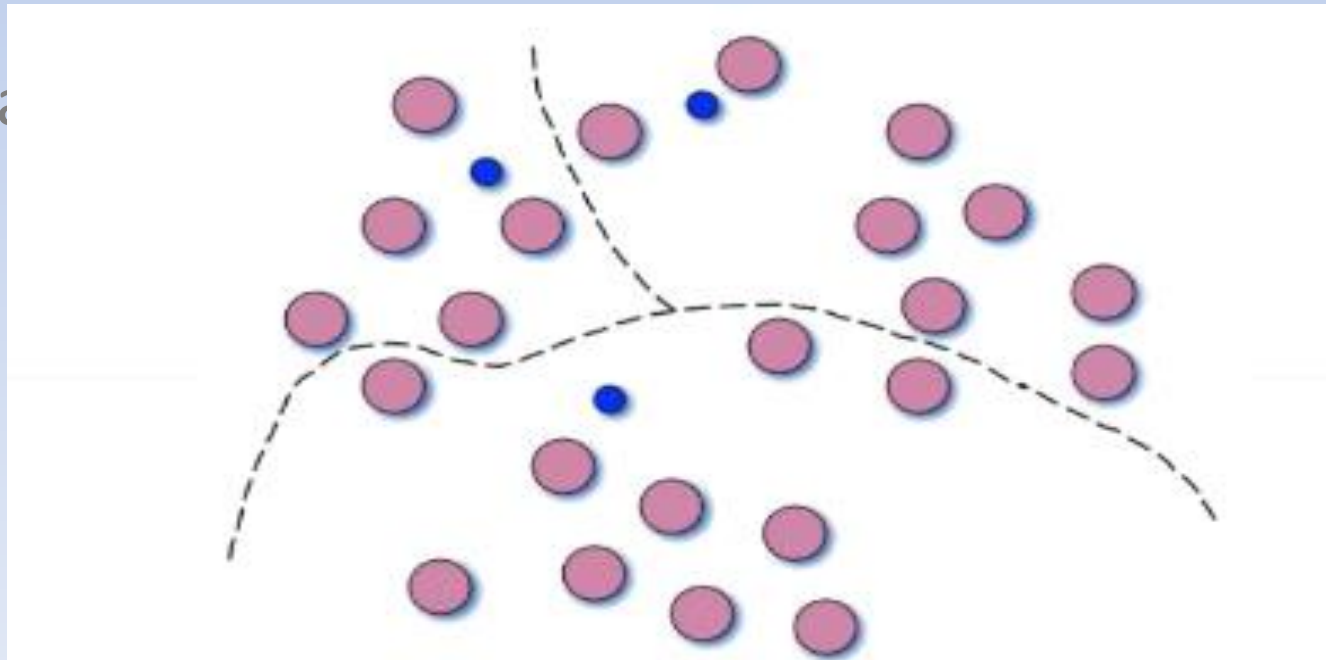
# K-means

## REPETER

- Affecter chaque point au cluster du centroïde le plus proche.

– Recalculer les centroïdes de chaque cluster.

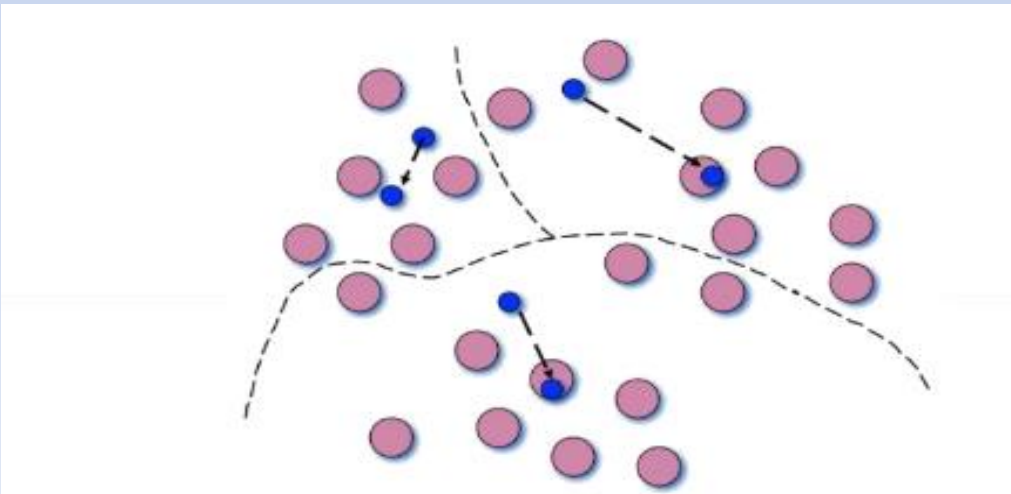
JUSQU'À



# K-means

## REPETER

- Affecter chaque point au cluster du centroïde le plus proche.
- Recalculer le centre de chaque cluster.

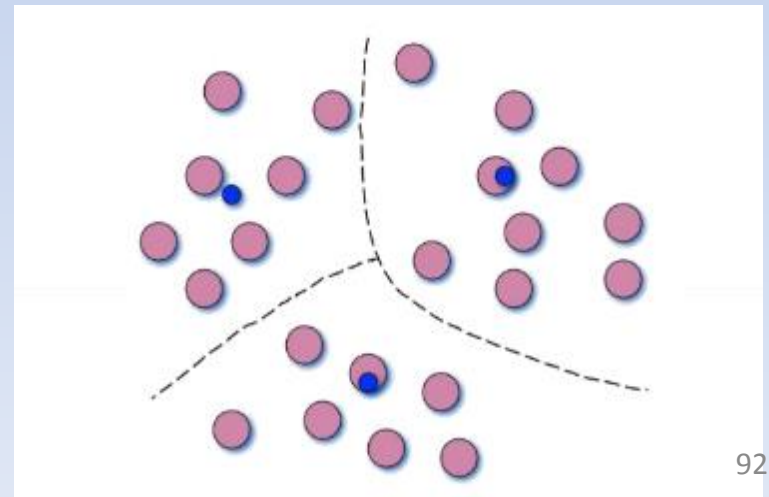


# K-means

## REPETER

- Affecter chaque point au cluster du centroïde le plus proche.
- Recalculer le centre de chaque cluster.

**JUSQU'À CONVERGENCE**



# K-means

## ***Remarque:***

La convergence correspond au fait que les centroïdes ne changent pas après une mise à jour.

***Attention*** : La convergence des centroïdes n'est pas garantie dans cet algorithme. Il faut en tenir compte dans lors de l'implémentation et ajouter une autre condition de sortie pour la boucle principale.



# K-means

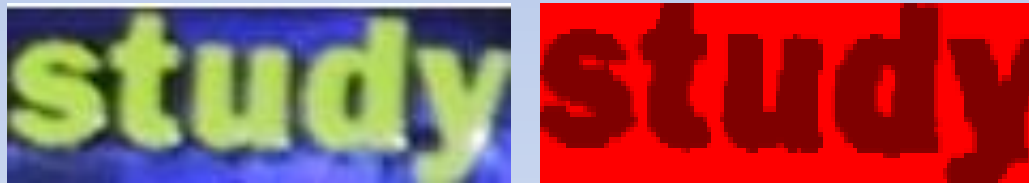
## *Remarque:*

La convergence correspond au fait que les centroïdes ne changent pas après une mise à jour.

**Attention :** La convergence des centroïdes n'est pas garantie dans cet algorithme. Il faut en tenir compte dans lors de l'implémentation et ajouter une autre condition de sortie pour la boucle principale.

# K-means

- K-means pour la segmentation d'images



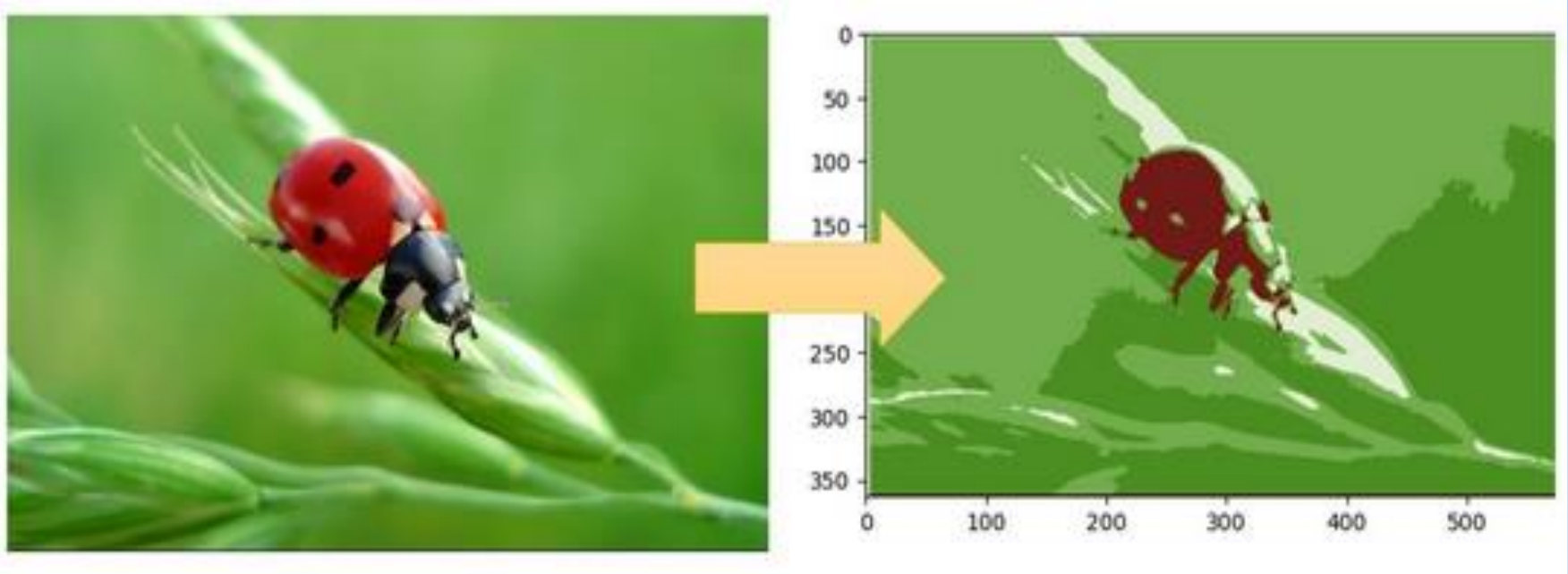
# K-means

- K-means pour la segmentation d'images



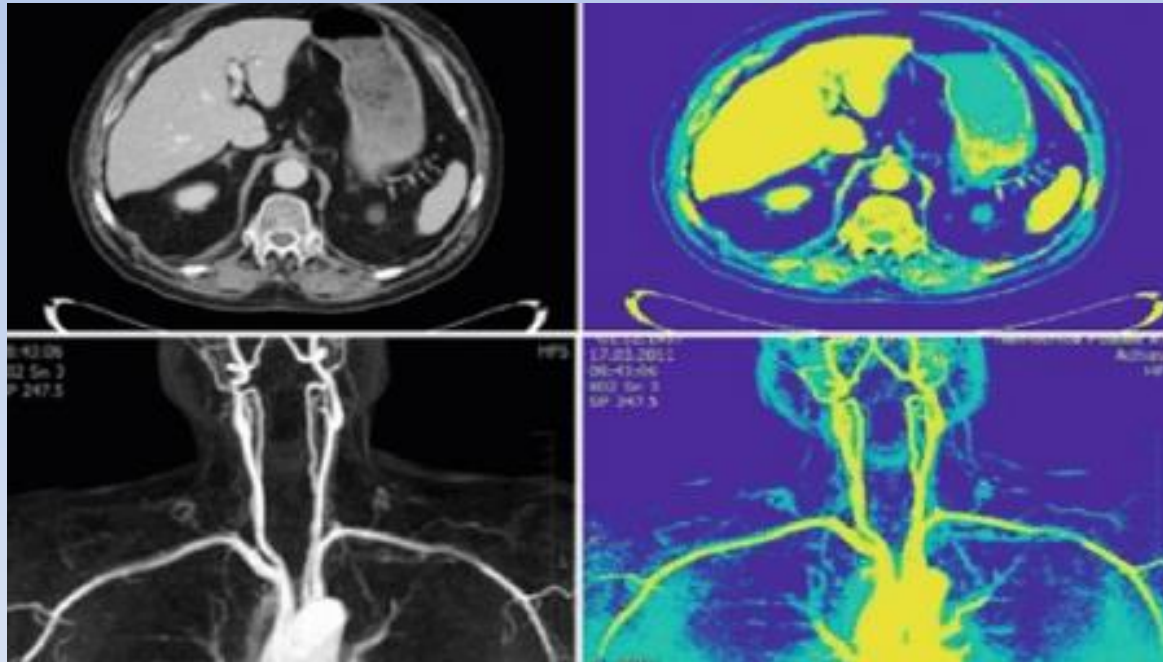
# K-means

- K-means pour la segmentation d'images



# K-means

- K-means pour la segmentation d'images médicales.



# Avantages du k-means

# Avantages du k-means

L'algorithme de k-means est:

- 1) Très facile à comprendre et à mettre en œuvre.
- 2) Simple et rapide.
- 3) Applicable à des données de grandes tailles, et aussi à tout type de données (mêmes textuelles), en choisissant une bonne notion de distance.

# Avantages du k-means

L'algorithme de k-means est:

- 1) Très facile à comprendre et à mettre en œuvre.
- 2) Simple et rapide.
- 3) Applicable à des données de grandes tailles, et aussi à tout type de données (mêmes textuelles), en choisissant une bonne notion de distance.



# Avantages du k-means

L'algorithme de k-means est:

- 1) Très facile à comprendre et à mettre en œuvre.
- 2) Simple et rapide.
- 3) Applicable à des données de grandes tailles, et aussi à tout type de données (mêmes textuelles), en choisissant une bonne notion de distance.

# Inconvénients du k-means

# Inconvénients du k-means

- 1) Le nombre de classe doit être fixé au départ.
- 2) Le résultat dépend de l'initialisation des centres des classes.
- 3) Les clusters sont construits par rapports à des objets inexistantes (les milieux)

# Inconvénients du k-means

- 1) Le nombre de classe doit être fixé au départ.
- 2) Le résultat dépend de l'initialisation des centres des classes.
- 3) Les clusters sont construits par rapports à des objets inexistant (les milieux)

# Inconvénients du k-means

- 1) Le nombre de classe doit être fixé au départ.
- 2) Le résultat dépend de l'initialisation des centres des classes.
- 3) Les clusters sont construits par rapports à des objets inexistantes (les milieux)

# Types d'apprentissage automatique

Les méthodes d'apprentissage peuvent être classées en trois principales catégories:

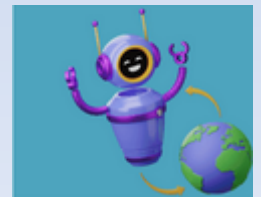
**Apprentissage  
supervisé**



**Apprentissage non-  
supervisé**



**Apprentissage par  
renforcement**



# Apprentissage par renforcement

L'apprentissage par renforcement  
(reinforcement learning) est:

- Un processus
- Dans lequel un agent (robot,...)
- Apprend à prendre des décisions
- A partir d'expérimentations et d'erreurs.

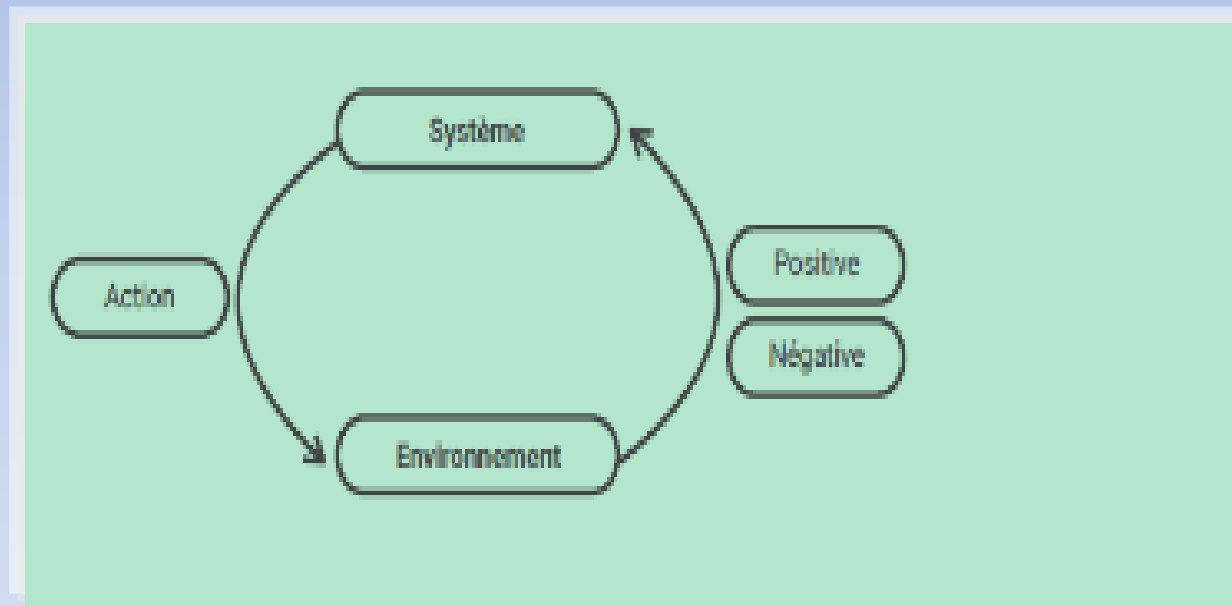
# Apprentissage par renforcement

L'apprentissage par renforcement est:

- Une technique de machine learning (ML) qui entraîne les logiciels à **prendre des décisions** en vue d'obtenir les **meilleurs résultats**.



# Apprentissage par renforcement



# Processus du ML/DL

# Processus du ML

**1. Collection des données:** Il s'agit de regrouper les données d'un problème à résoudre.

=> La construction du Dataset.

# Processus du ML

**2. Prétraitement des données:** Afin de rendre la Dataset utilisable à l'apprentissage, il faut le nettoyer:

- La suppression de données inutiles.
- La suppression de données répétées.
- La suppression de données incomplètes et manquantes.
- L'enrichissement par d'autres données, décomposition des données.

# Processus du ML

**2. Prétraitement des données:** Afin de rendre la Dataset utilisable à l'apprentissage, il faut le nettoyer:

- La suppression de données inutiles.
- La suppression de données répétées.
- La suppression de données incomplètes et manquantes.
- L'enrichissement par d'autres données, décomposition des données.

# Processus du ML

**2. Prétraitement des données:** Afin de rendre la Dataset utilisable à l'apprentissage, il faut le nettoyer:

- La suppression de données inutiles.
- La suppression de données répétées.
- La suppression de données incomplètes et manquantes.
- L'enrichissement par d'autres données, décomposition des données.

# Processus du ML

**2. Prétraitement des données:** Afin de rendre la Dataset utilisable à l'apprentissage, il faut le nettoyer:

- La suppression de données inutiles.
- La suppression de données répétées.
- La suppression de données incomplètes et manquantes.
- L'enrichissement par d'autres données, décomposition des données.

# Processus du ML

## 3. Choix du modèle:

Selon le problème traité, on peut choisir:

- La régression: s'il s'agit d'un problème de prédiction.
- Le clustering: pour les problèmes tels que la détection d'anomalies, la segmentation d'images, etc.
- Naïve bayes: s'il s'agit d'un problème de classification.



# Processus du ML

## 3. Choix du modèle:

Selon le problème traité, on peut choisir:

- La régression: s'il s'agit d'un problème de prédiction.
- Le clustering: pour les problèmes tels que la détection d'anomalies, la segmentation d'images, etc.
- Naïve bayes: s'il s'agit d'un problème de classification.

# Processus du ML

## 3. Choix du modèle:

Selon le problème traité, on peut choisir:

- La régression: s'il s'agit d'un problème de prédiction.
- Le clustering: pour les problèmes tels que la détection d'anomalies, la segmentation d'images, etc.
- Naïve bayes: s'il s'agit d'un problème de classification.

# Processus du ML

## 4. Entrainement:

Les données de Dataset sont séparées en:

- 80% pour entraîner l'algorithme choisi.
- 20% pour tester et vérifier la performance du résultat.

# Processus du ML

**5. Evaluation:** L'évaluation (l'étude des valeurs prédictives) permet:

- De définir si le modèle du machine learning est fiable.
- Dans quels cas il commet des erreurs.
- Dans quelle mesure.

Merci pour votre attention