

- **Contenu de la séance**
 - Méthodes d'apprentissage
 - supervisé
 - par renforcement
 - non supervisé
 - Outils : Représentation, paramètres, critères

Méthodes d'apprentissage

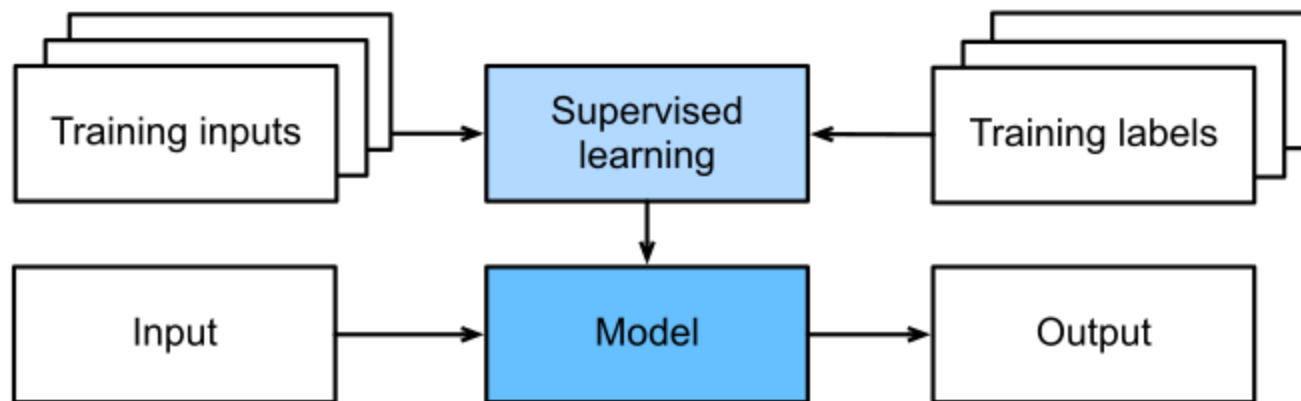
- Trois grandes techniques (bien connues) d'apprentissage : **Supervisé, par renforcement** et non supervisé

App. Supervisé :

Objectif : prédire à partir d'entrées, le résultat est un modèle

Prédire quoi ? Labels, valeurs, ...

Entrées : caractéristiques



App. supervisé

- Types de prédictions :

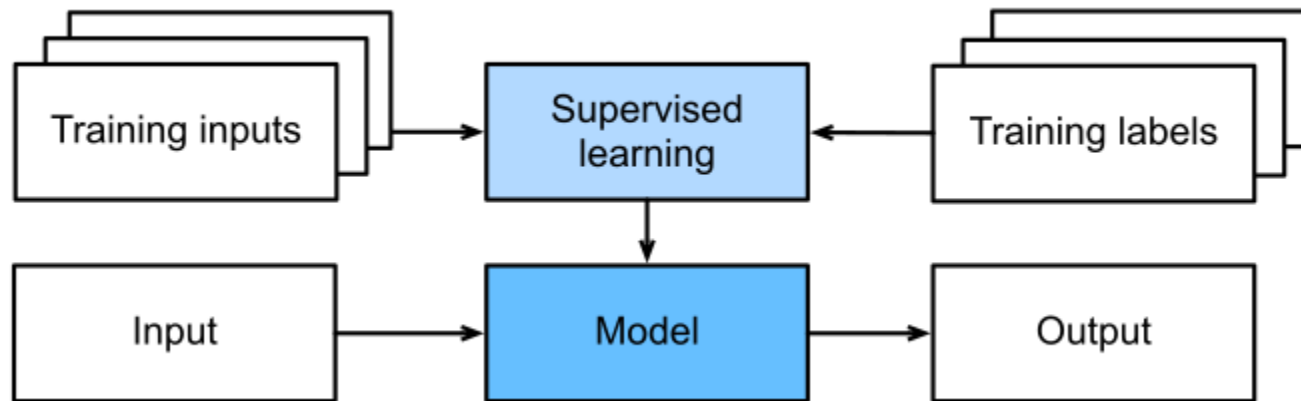
- **Régression**

Ex. Combien de temps est nécessaire pour faire l'itinéraire Paris-Nantes ?

Combien de chiens apparaissent dans une photo ?

- L'objectif du processus d'apprentissage est celui de minimiser l'écart entre les prédictions et les valeurs attendues
- Critères :

$$\sum_i |y_i - y'_i| \quad \sum_i (y_i - y'_i)^2$$



App. supervisé

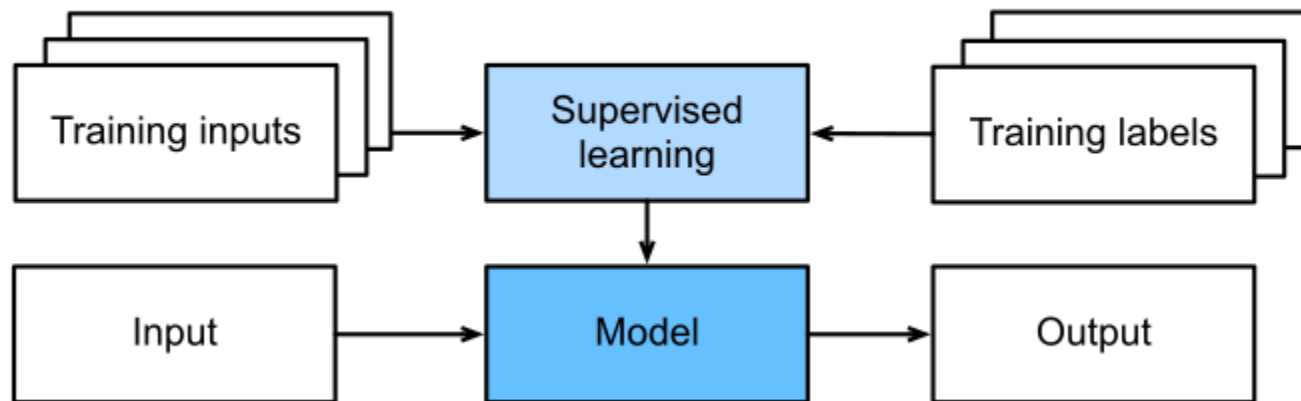
- Types de prédictions :

- **Classification**

Ex. Reconnaissance de texte / vocale

Classification binaire / multi-classes

- L'objectif du processus d'apprentissage est celui d'optimiser le modèle afin d'effectuer une assignation de classes
- Critère : entropie-croisée (critère probabiliste)



App. supervisé

– Autres types de prédictions :

– **Labélisation**

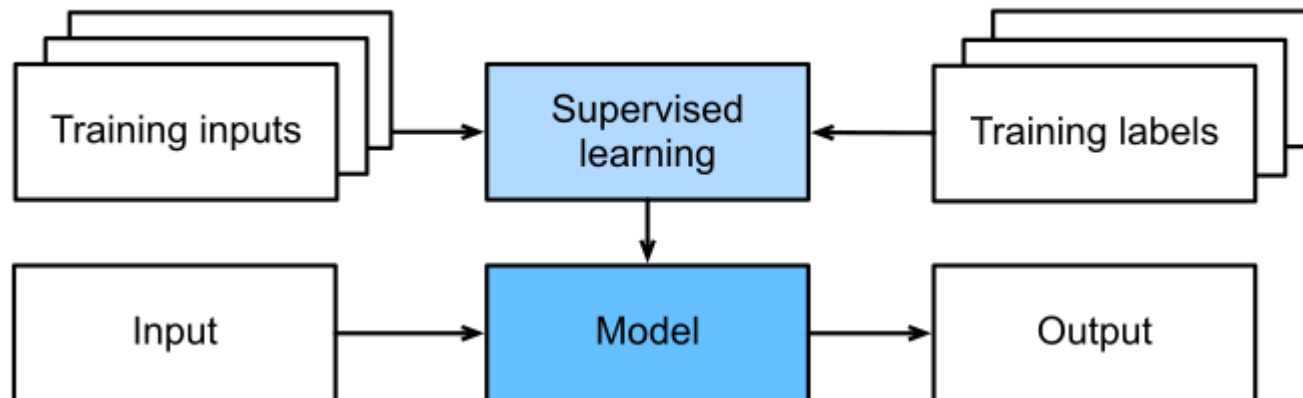
Ex. Classification où les classes prédites ne sont pas exclusives

– Recherche et classement

Ex. Classement des résultats de recherche les plus pertinents

– Systèmes de recommandation

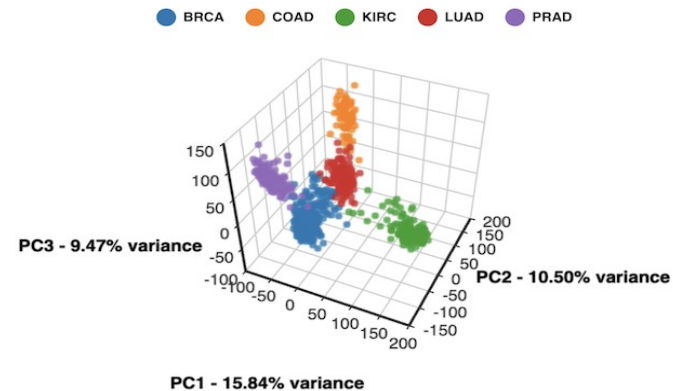
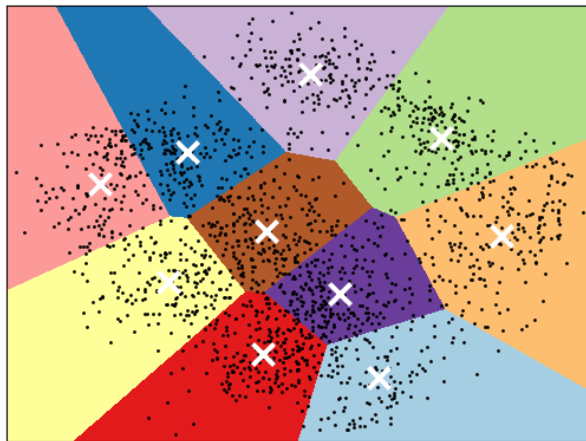
Ex. Prédiction d'objets à acheter



App. non supervisé

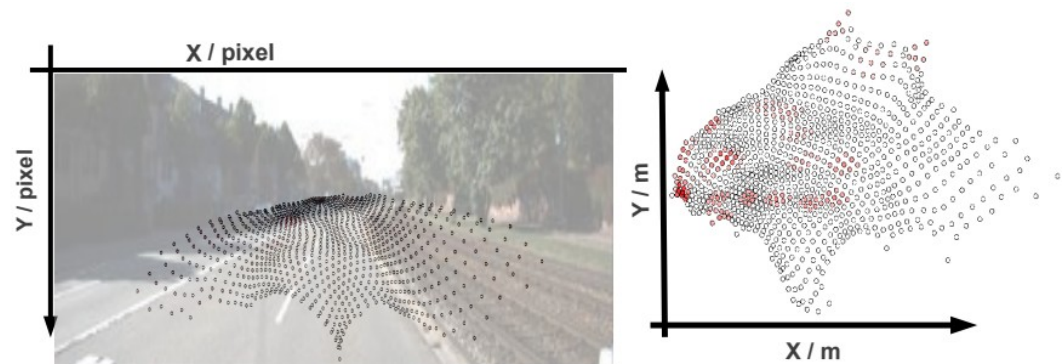
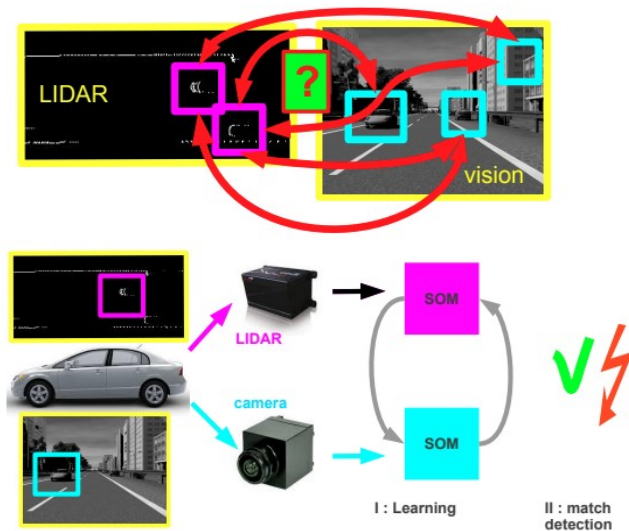
- Méthodes capables d'inférer à partir des données :
- Clustering : Un regroupement de données partageant une ou plusieurs caractéristiques
- Analyse en Composantes Principales : les caractéristiques essentielles à départager les données en groupes
- Apprentissage de représentations : établir de relations entre les données
- Cartes auto-organisatrices (Self-Organizing Maps)

K-means clustering on the digits dataset (PCA-reduced data)
Centroids are marked with white cross



App. non supervisé

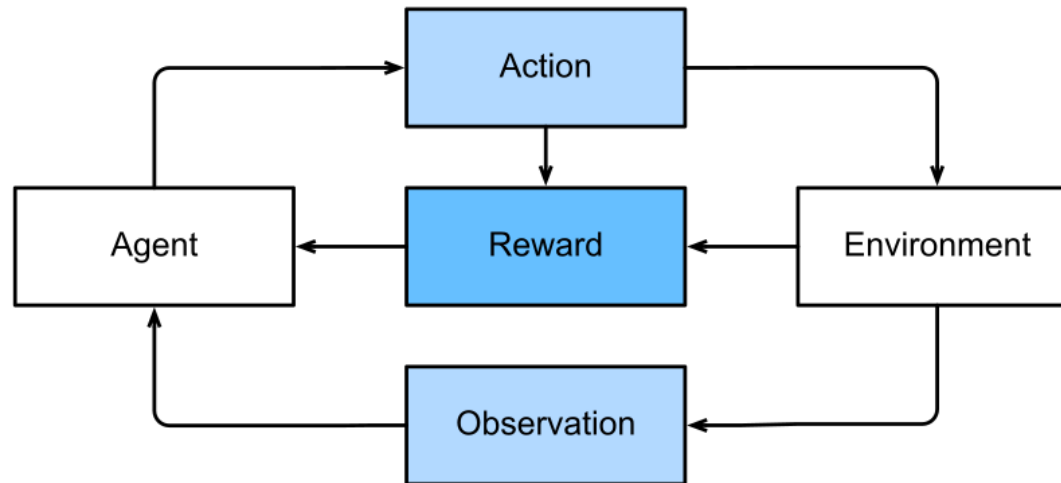
- Méthodes capables d'inférer à partir des données :
- Clustering : Un regroupement de données partageant une ou plusieurs caractéristiques
- Analyse en Composantes Principales : les caractéristiques essentielles à départager les données en groupes
- Apprentissage de représentations : établir de relations entre les données
- Cartes auto-organisatrices (Self-Organizing Maps)



Visualiser des données complexes dans un espace bidimensionnel (la carte) tout en préservant les relations topologiques entre les données.

App. par renforcement

- Objectif : déterminer une bonne « politique » (policy) d'échange entre les observations et les actions
- Critère : défini par une stratégie de récompense
- Ex. Conduire une voiture
- Nécessité d'un environnement pour déterminer si les actions sont adéquates.



Outils : Représentation, paramètres, critères

- Bibliothèque **Scikit-Learn**
- Trois instances seront abordées :

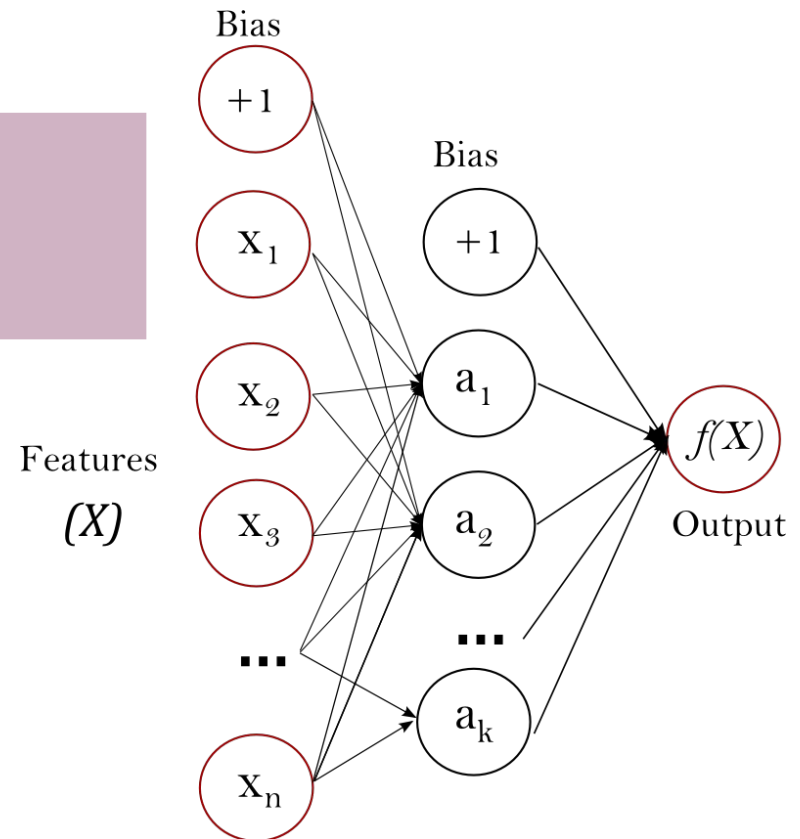
Perceptron (un seul neurone)

Perceptron multicouche (MLP)

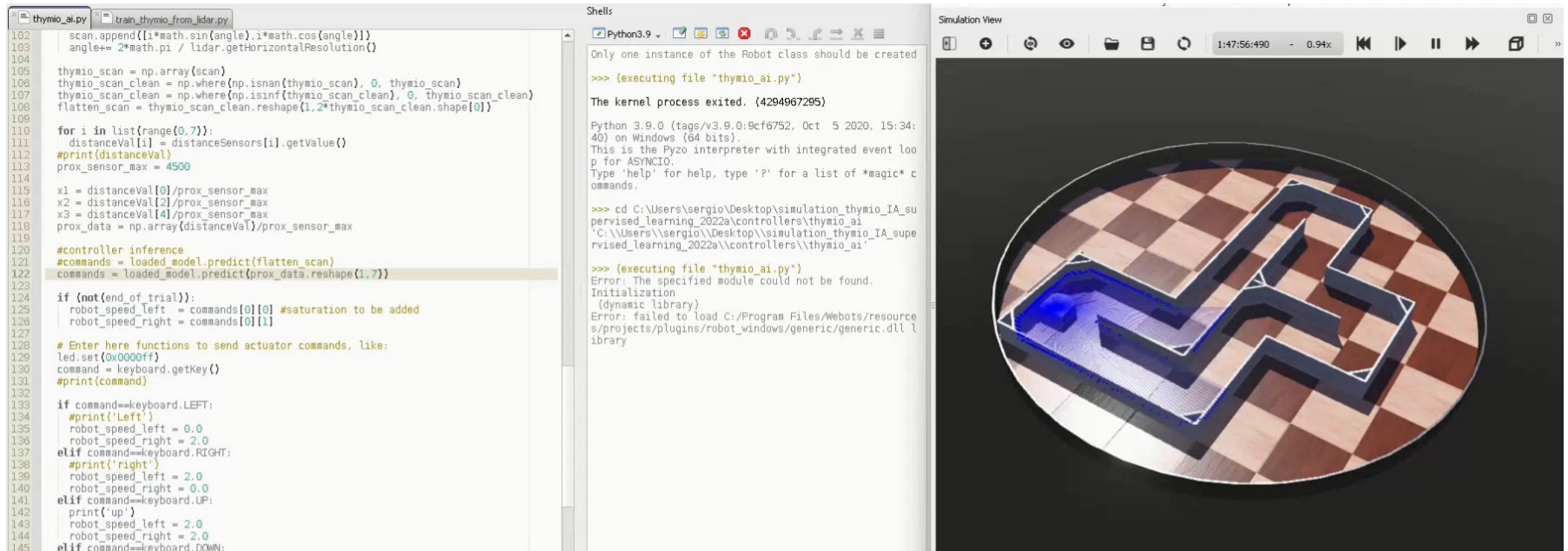
Classifieur (Classification binaire)

- **Regressor (Régression)**

- Méthodes :
 - Fit : apprentissage
 - Coeffs_ : poids
 - Intercepts_ : biais
 - Activation : fonction d'activation
 - out_activation : fonction d'activation en sortie
 - Score : précision moyenne (classif : match/samples) / (regr : R^2)



Résultats attendus : application



- Dataset : proximateurs et consignes de vitesse
- Modèle : Script faisant la lecture du modèle et le test avec le dataset
- Vidéo de validation