



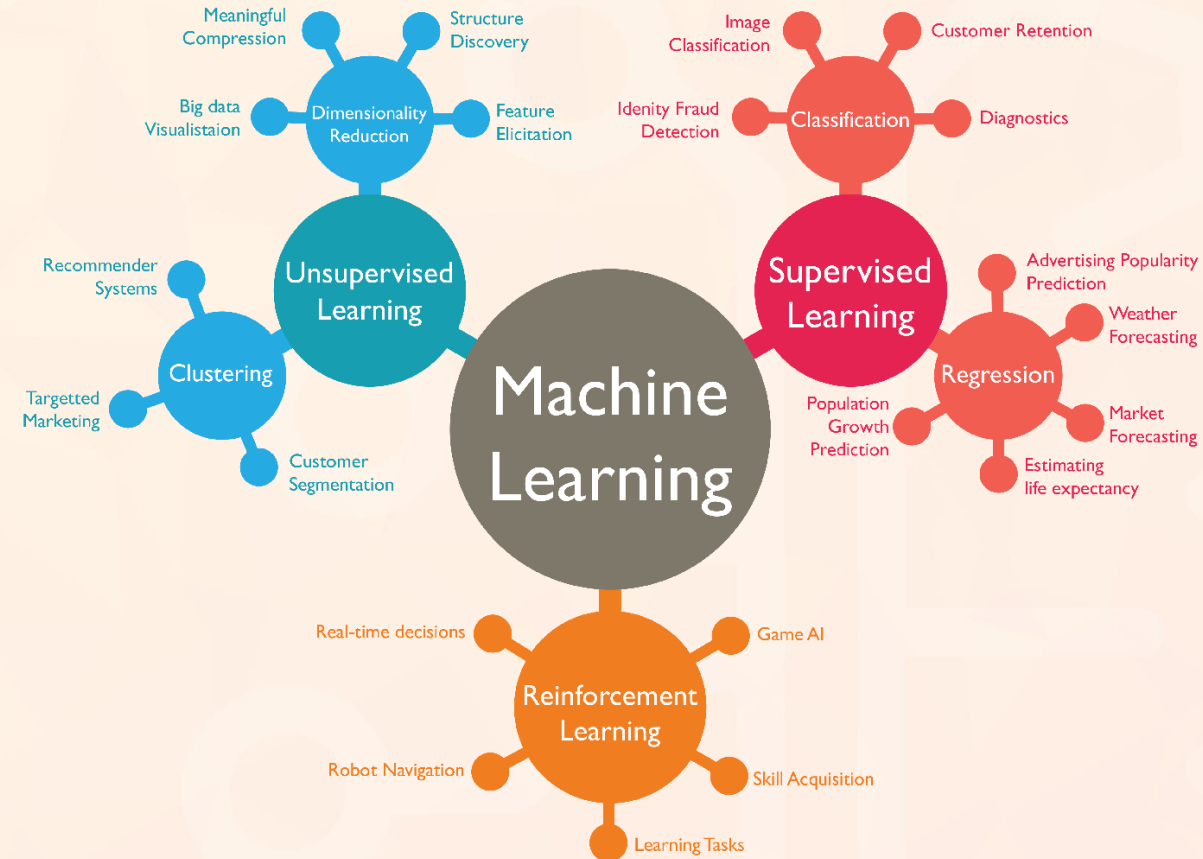
Découverte des outils de Machine Learning

Projet de permanence

Introduction



- **Machine Learning** : domaine vaste avec de nombreux algorithmes.
- Bibliothèques regroupant des implémentations de ces algorithmes => Ne pas réinventer la roue à chaque fois !
- Connaissance nécessaire pour faire du ML en pratique.



Objectifs de ce sujet



- Introduction de deux bibliothèques fréquemment utilisées en Machine Learning (Keras, et scikit-learn).
- Présentation de nombreuses techniques.
- Et vous permettre d'avoir vos premiers résultats de ML !

Plan



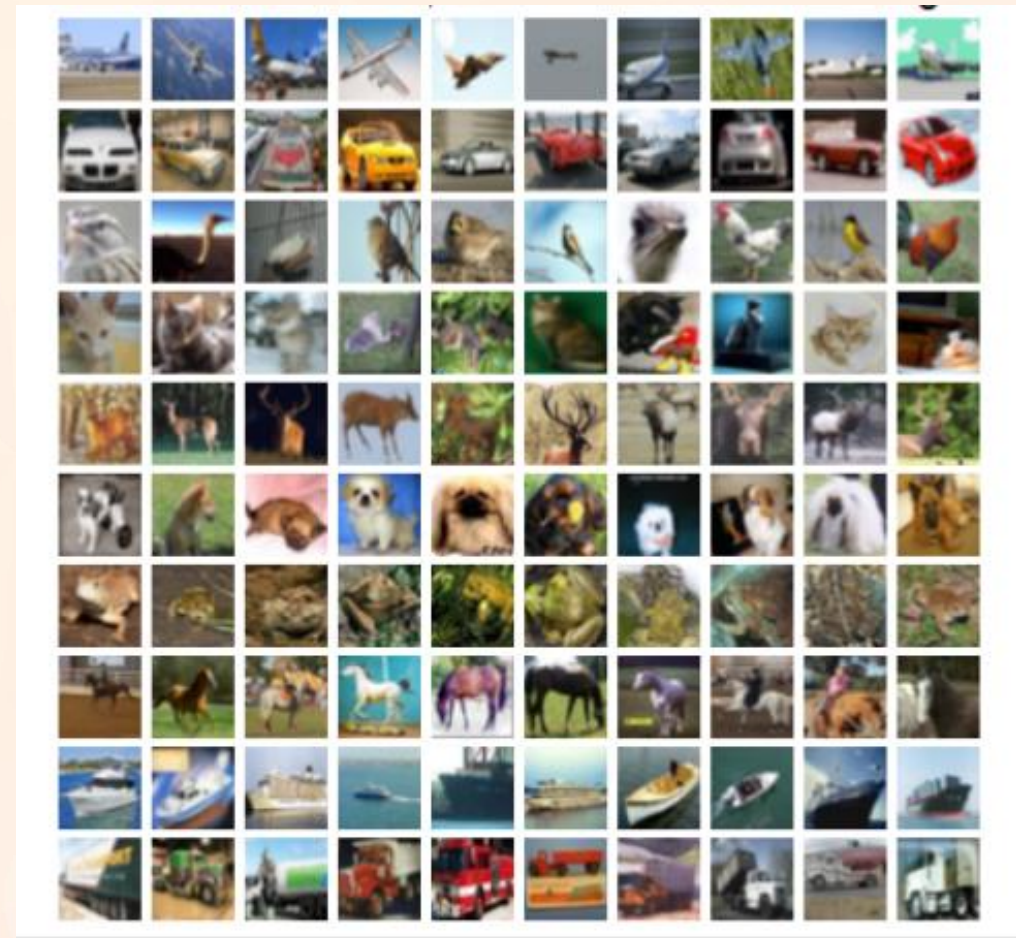
1. Classification supervisée à l'aide de réseau de neurones.
2. Les auto-encodeurs, et applications diverses.
3. Classification non-supervisée sans réseau de neurones.

Présentation rapide



Utilisation du dataset CIFAR-10 :

- Des images de 32 par 32 en couleur
- 10 classes
- 5000 images d'entraînement par classe
- 1000 images de test par classe



Présentation rapide



TensorFlow :

- Librairie développée par Google pour créer et manipuler des NN facilement.
- Existe d'autres librairies dans le même style : PyTorch (Facebook) ou Caffe (Berkley).

Keras :

- Sur-couche de TensorFlow qui facilite encore plus la construction de réseaux de neurones.
- Librairie pour rapidement construire divers types de réseaux de neurones et les entraîner.
- Inclus quelques datasets par défaut, ce qui est pratique pour s'entraîner ou essayer des algorithmes.

Scikit-learn :

- Nombreux outils pour faire de l'analyse prédictive de données (Classification, régression, clustering, réduction de dimension, etc.).

Avertissement



- Beaucoup de nouvelles notions : vous allez être perdus au départ, c'est normal.
- Introduction de beaucoup de concepts, sans présenter la théorie. C'est le côté pratique qui nous intéresse ici.
- Sujet long, donc pas de panique si vous ne le finissez pas en 3h.



Les fichiers Jupyter



Sujet sous la forme d'un fichier Jupyter :

- Soit vous avez directement tout ce qu'il faut sur votre ordinateur (peu probable).
- Soit vous utilisez Google colab' pour le lancer (plus simple).

Utilisation des fichiers Jupyter :

- Alternance de bloc de texte et de bloc de code.
- Pour exécuter un bloc, faites « **Crtl** » + **R** (reste sur le bloc courant) ou « **Shift** » + **R** (Passe au bloc suivant).
- Tous vos print et plot s'affichent en dessous du bloc de code.

Comment commencer ?



Vous commencez à connaître la routine :

1. Télécharger le sujet : <https://automatants.cs-campus.fr/projets/>
2. Chercher Google Colab' sur votre moteur de recherche favori.
3. Ouvrez Google Colab', et cliquez sur « Importez » sur la page qui s'ouvre.

Si avez des questions, n'hésitez pas à les poser à un 2A.