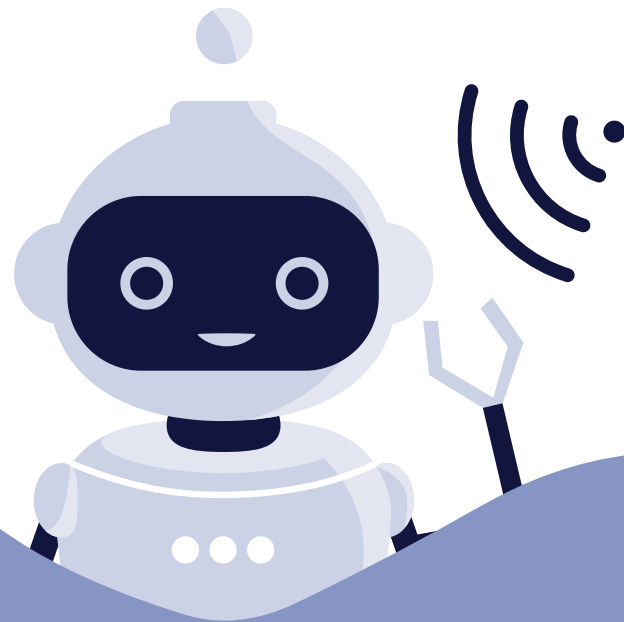
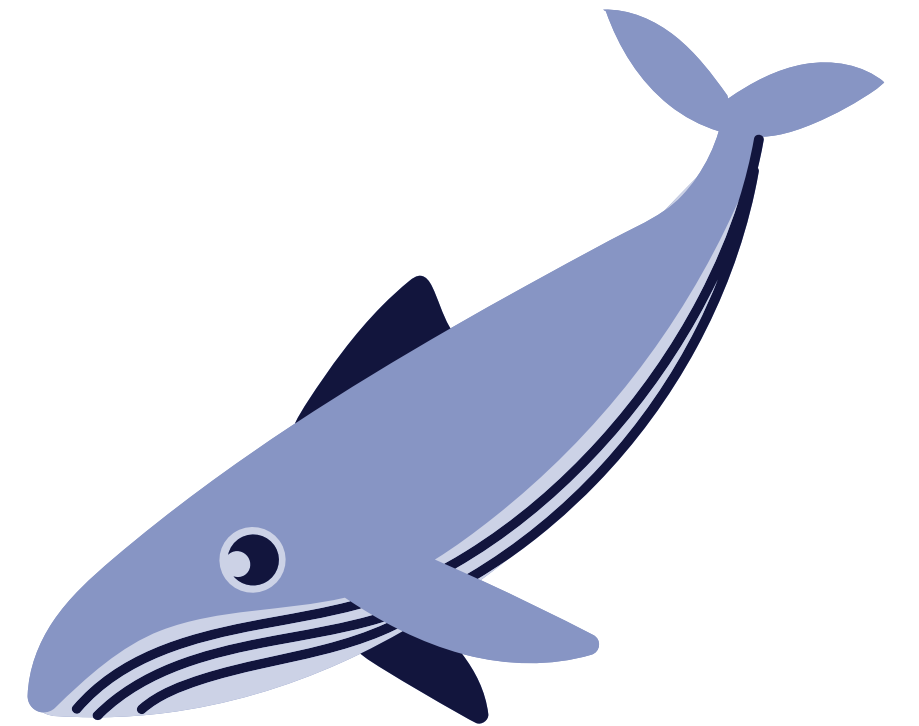
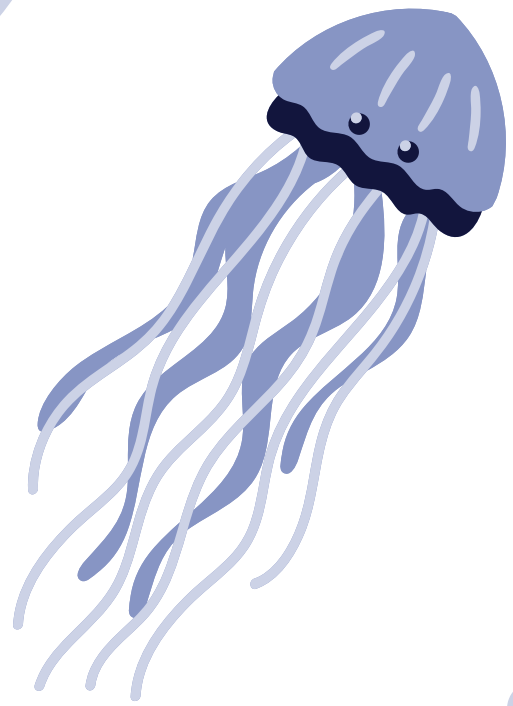


ANALYSER ET CLASSIFIER LES SONS DE MAMMIFÈRES MARINS AVEC LES OUTILS D'IA



Eléa Petton

PRÉSENTATION

 **OVHcloud**
Machine Learning Engineer

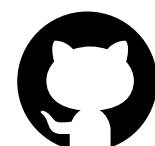
ISEN
ALL IS DIGITAL!
BREST

 yncréa

Etudiante ingénieure



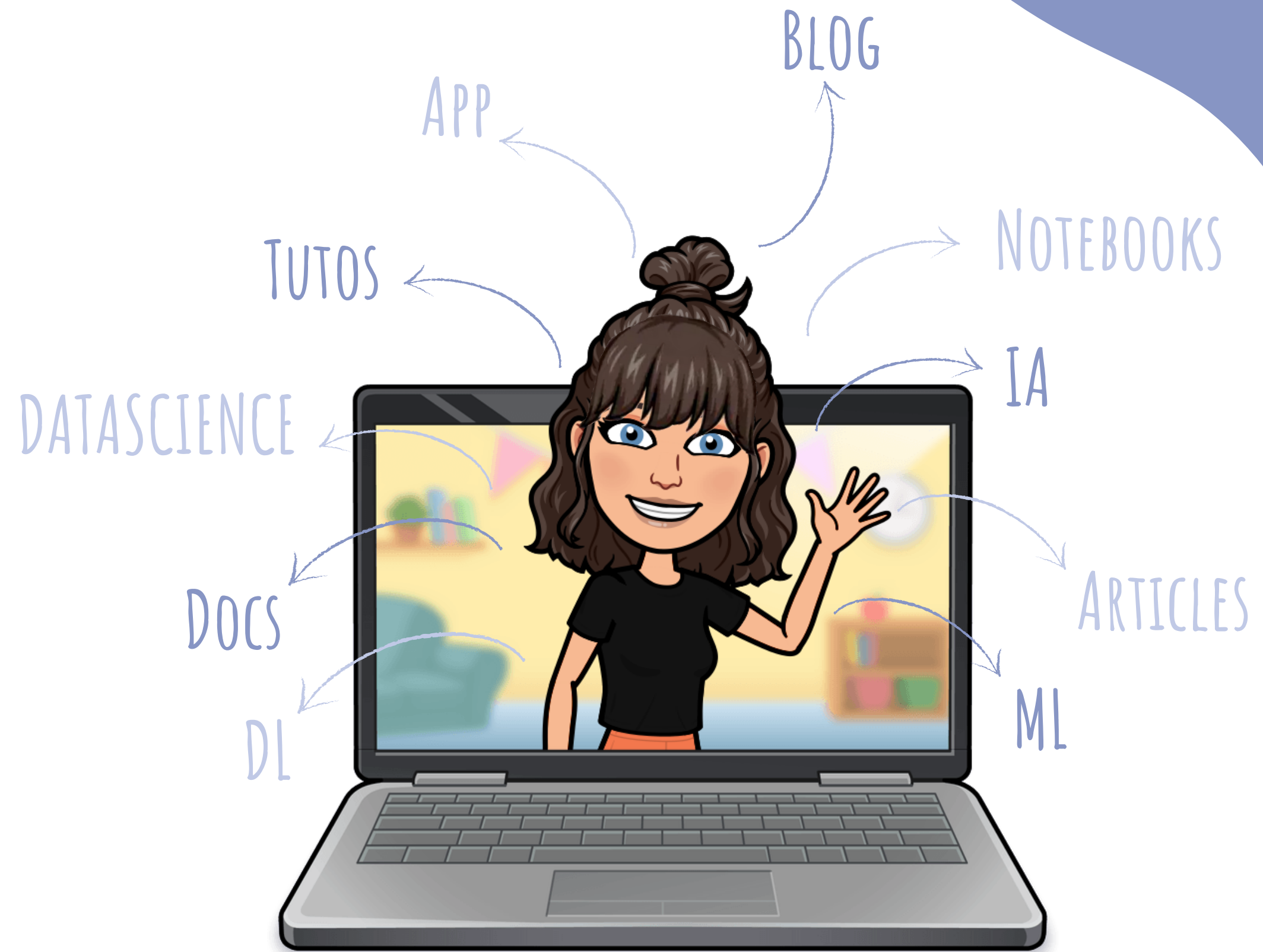
@EleaPetton



eleapttn



Eléa PETTON



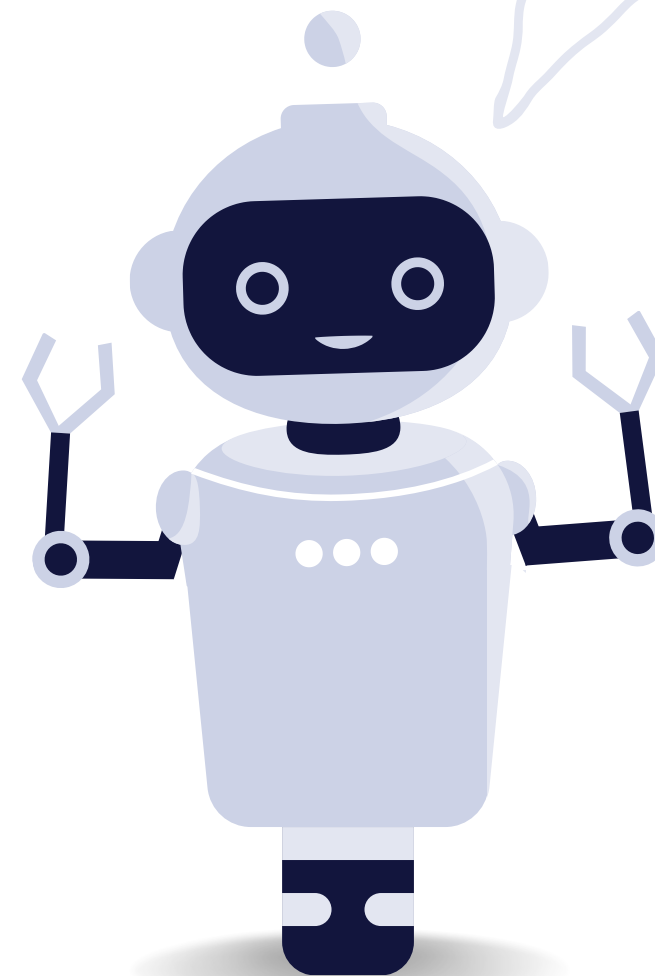


QU'EST-CE QUE L'IA ?



Toi humain tu sais
différencier le chien
du chat, mais pas moi.

Comment faire pour
l'apprendre ?

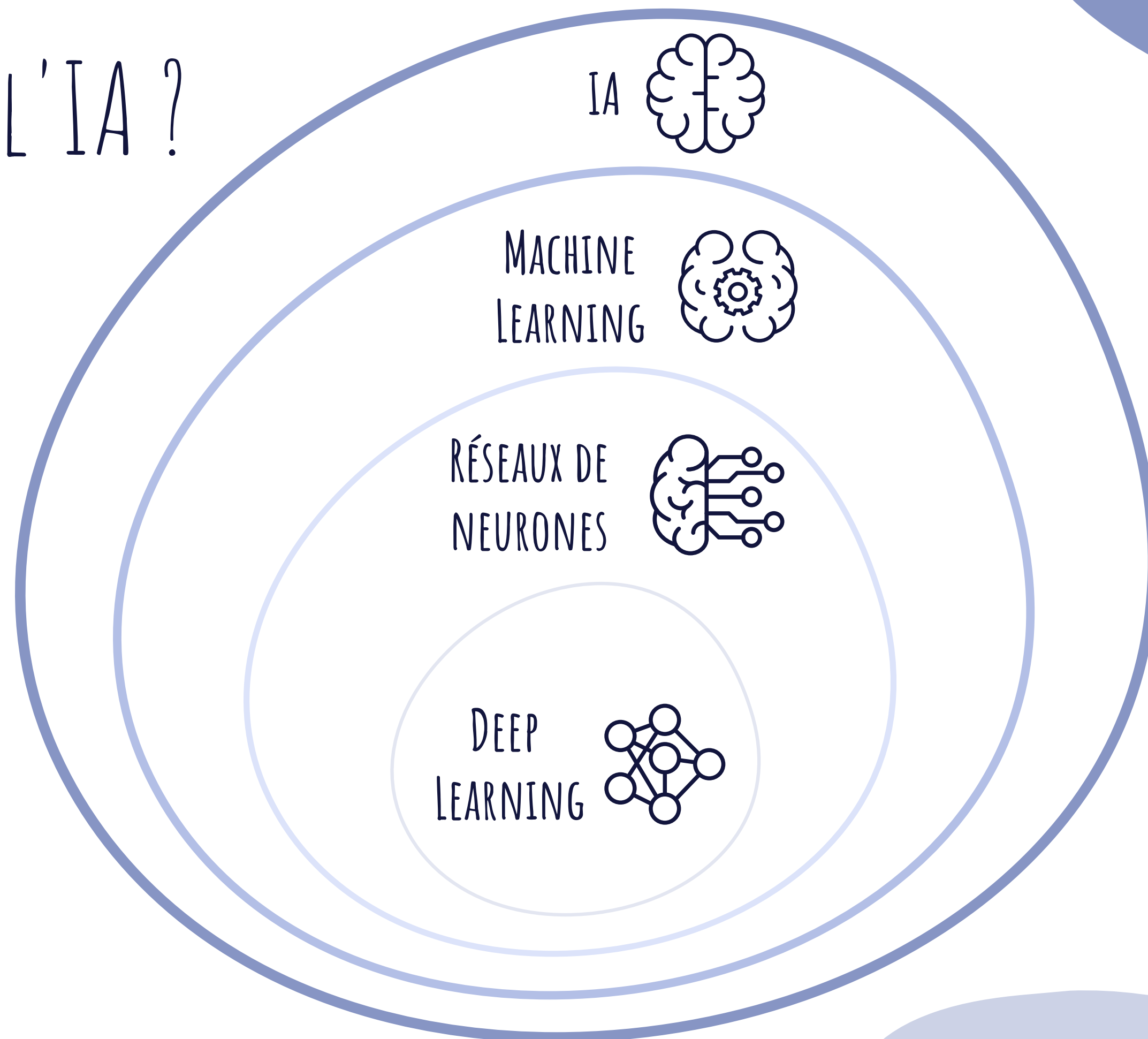
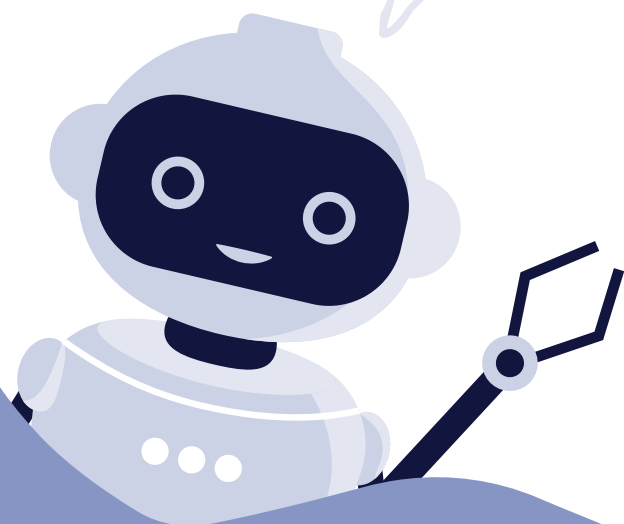


Attends, je vais te
donner un algorithme
qui va te permettre
d'apprendre.

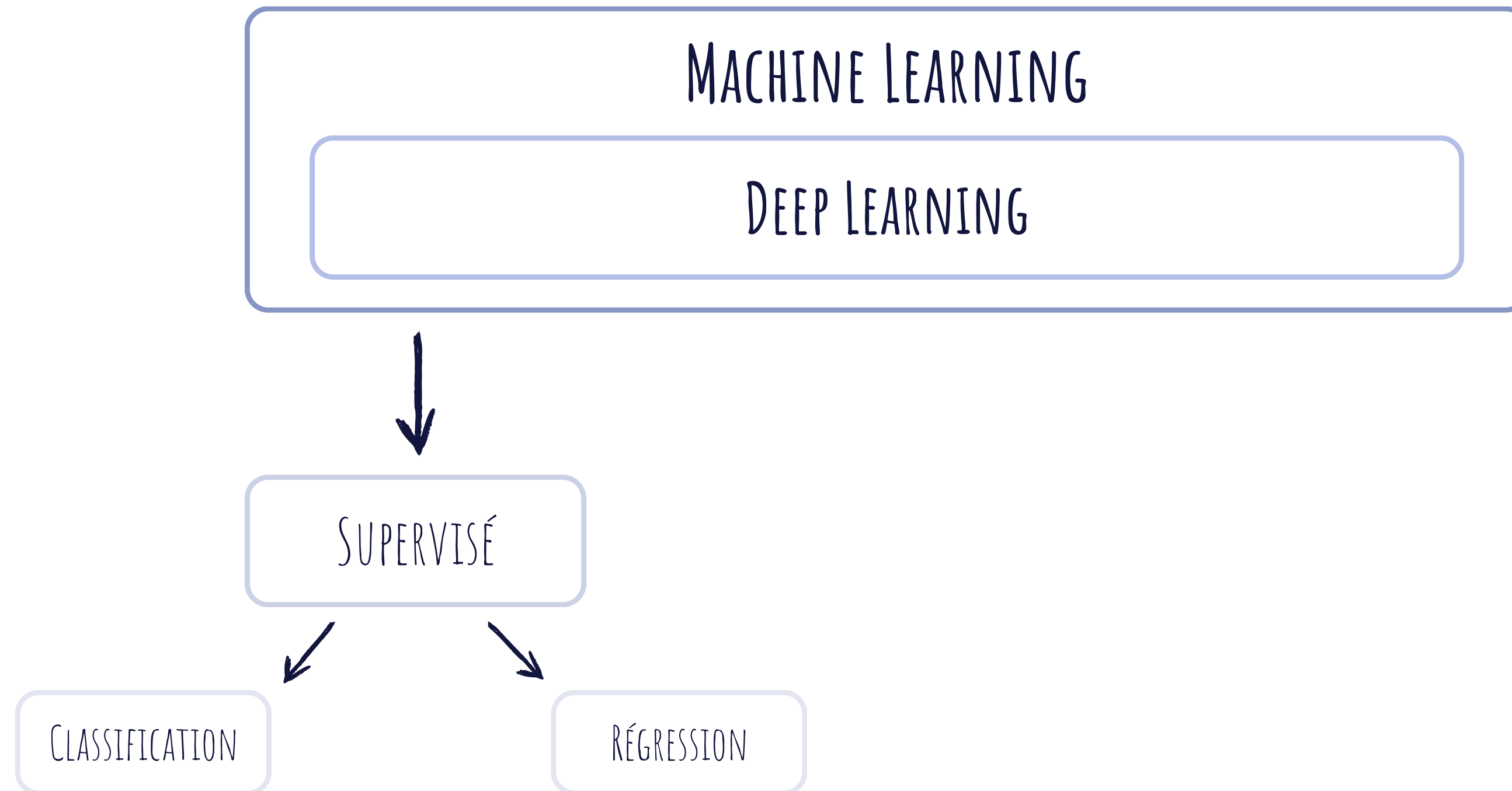


QU'EST-CE QUE L'IA ?

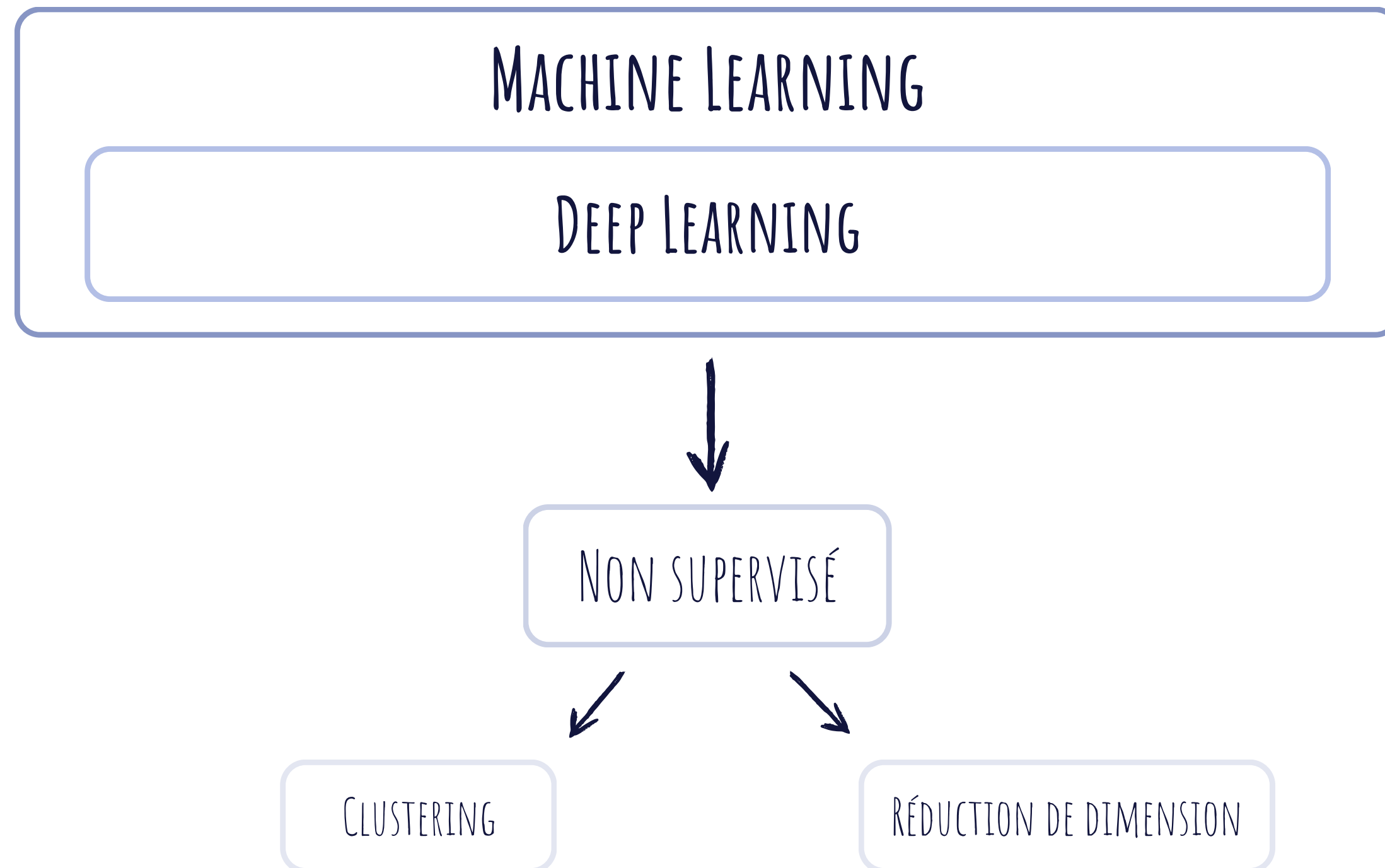
Je me demande
quelle est la
différence entre
tous ces termes ?



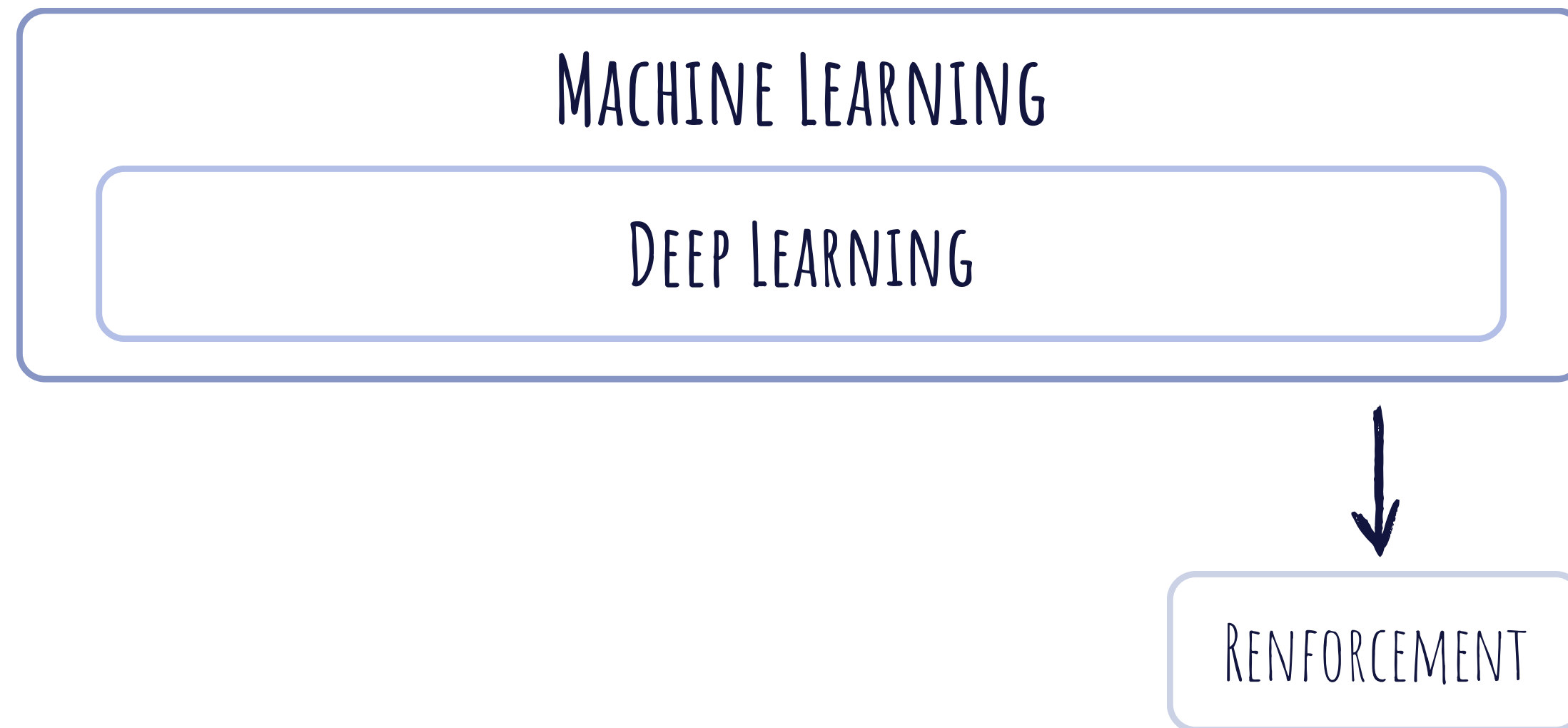
LE MACHINE LEARNING PLUS EN DÉTAIL



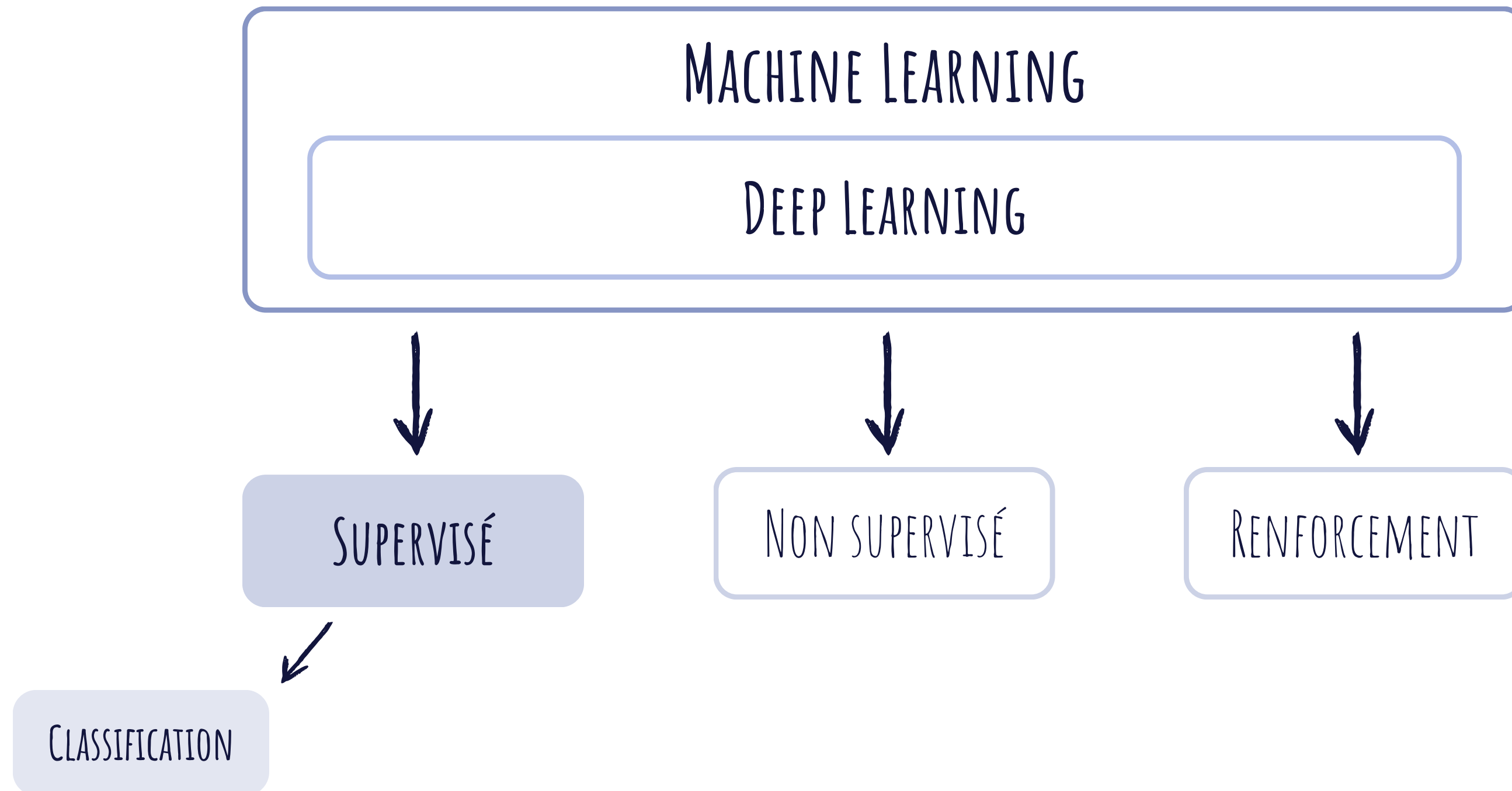
LE MACHINE LEARNING PLUS EN DÉTAIL



LE MACHINE LEARNING PLUS EN DÉTAIL



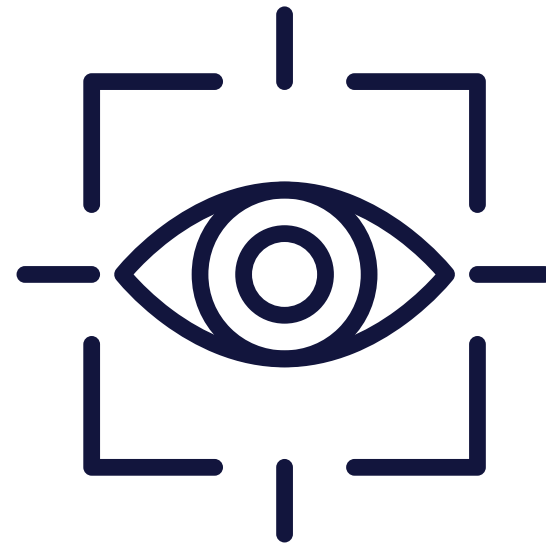
LE MACHINE LEARNING PLUS EN DÉTAIL



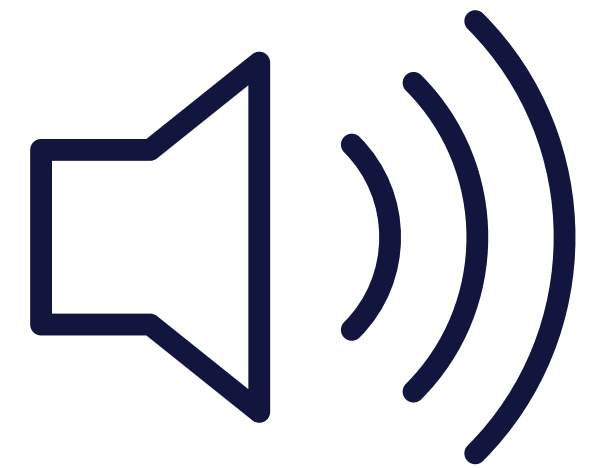
DIFFÉRENTES FAMILLES EN IA



NLP



Computer Vision



Audio

POURQUOI LE SON EN IA ?



Classification des données acoustiques



Classification des sons environnementaux



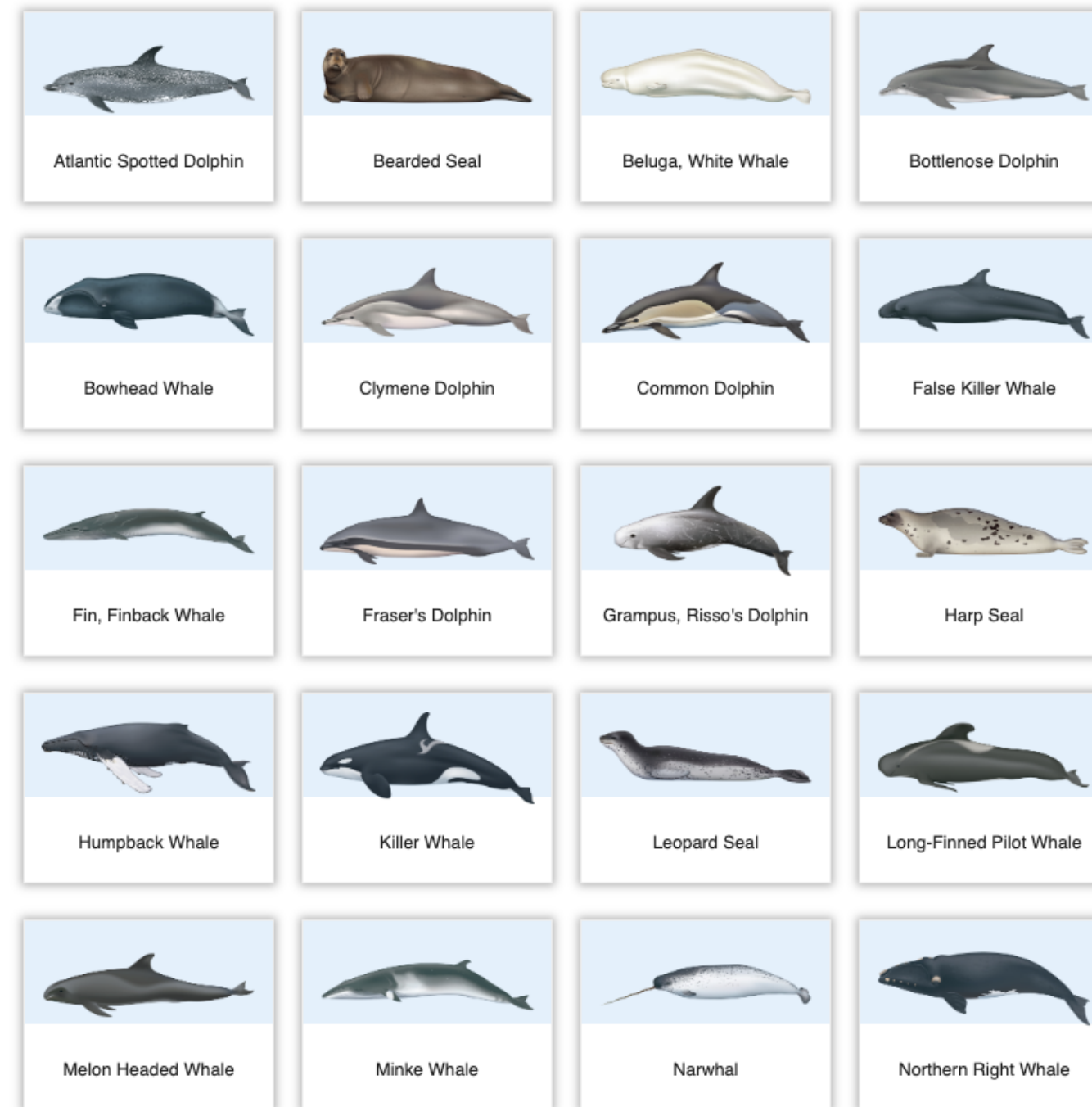
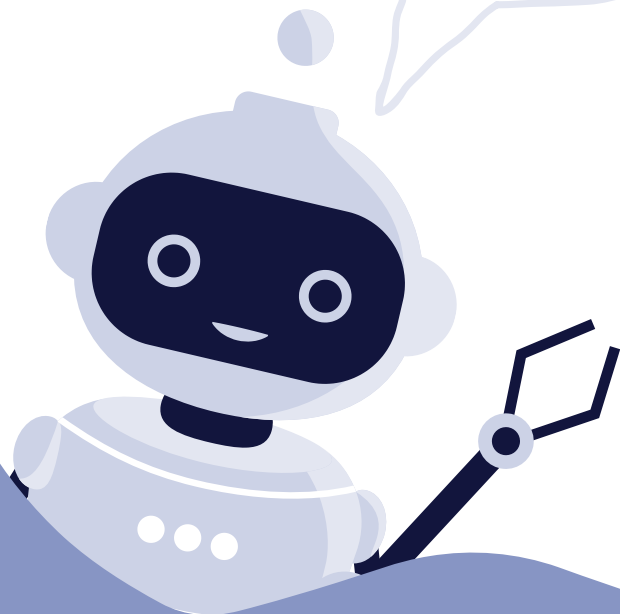
Classification du langage naturel



Classification de la musique

SONS DE MAMMIFÈRES MARINS

On va essayer de
différencier 45
mammifères marins
différents



Watkins Marine Mammal Sound Database

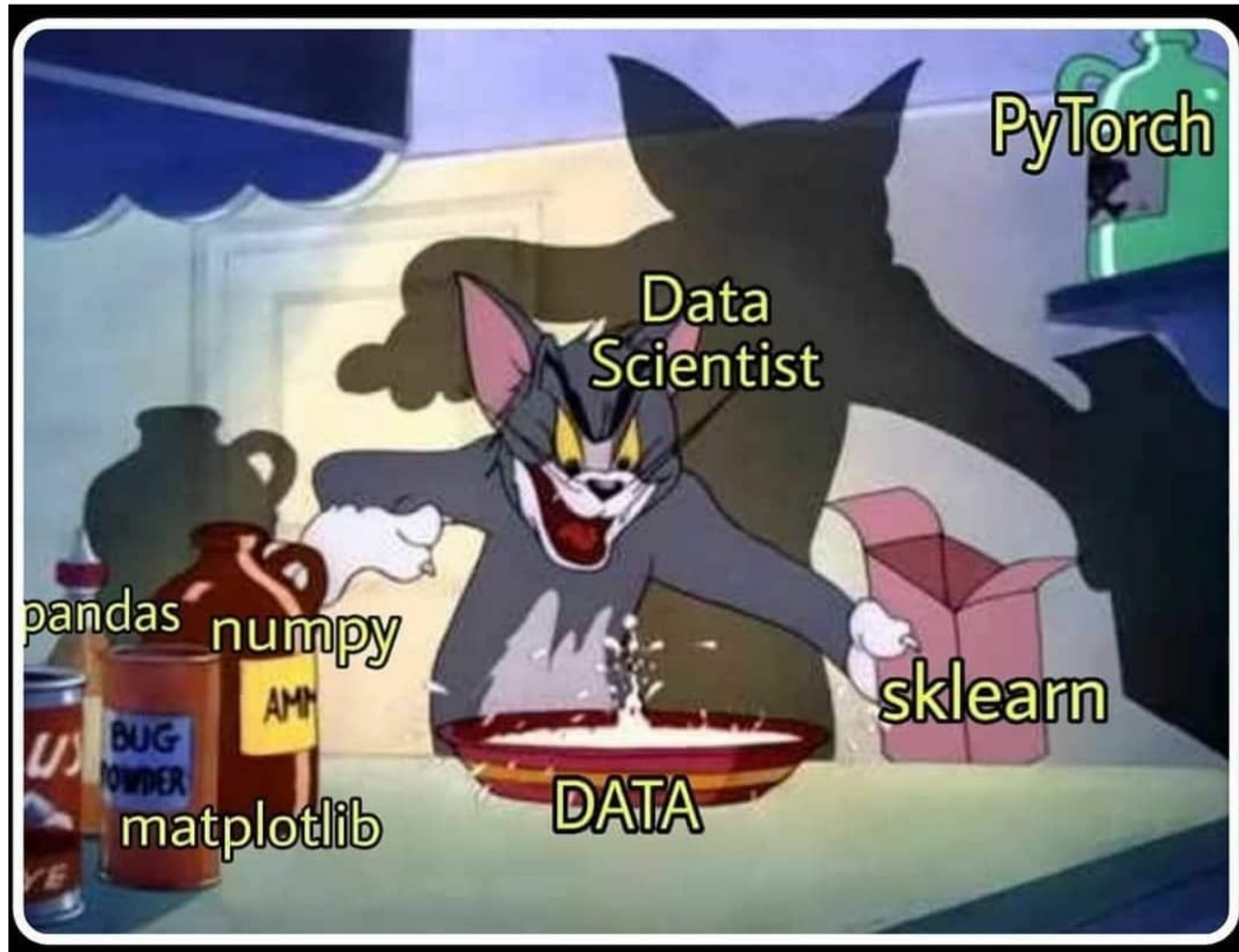
IMPORTANCE DU PRÉ-TRAITEMENT DES DONNÉES



QUALITÉ DES DONNÉES



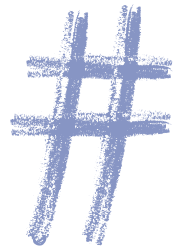
INFORMATIONS UTILES



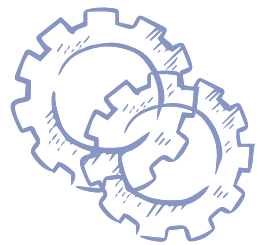
EXEMPLES DE PRÉ-TRAITEMENT POUR LES DONNÉES



Gestion des valeurs nulles



Standardisation



Gestion des variables qualitatives

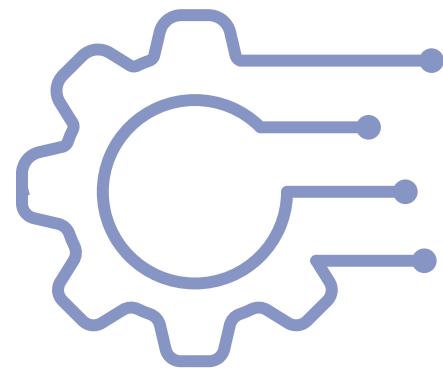


Encodage à chaud

ET POUR LES DONNÉES AUDIOS ?



audios .wav



paramètres audio
Librosa



26 paramètres audio Librosa



name	length	...	...	...	label

fichier .csv



**me cleaning
the data**

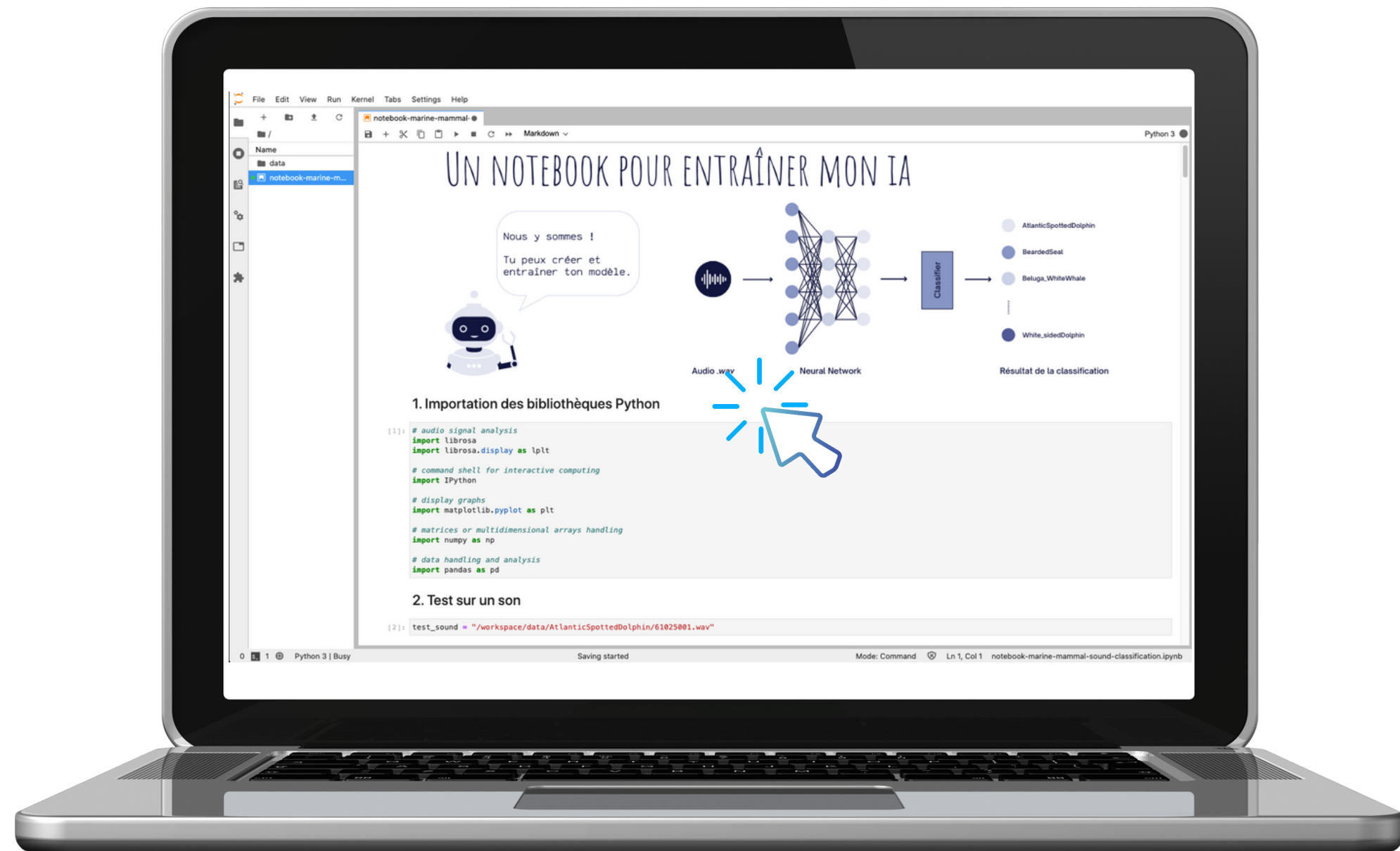
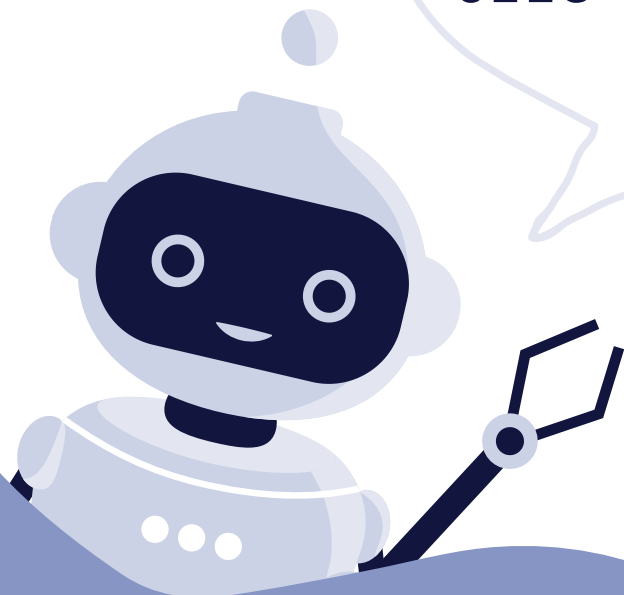


**me building
the model**

UN NOTEBOOK POUR ENTRAÎNER MON IA

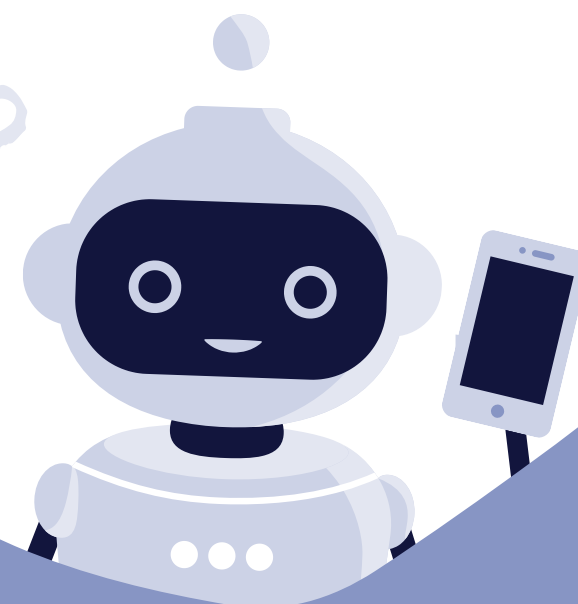
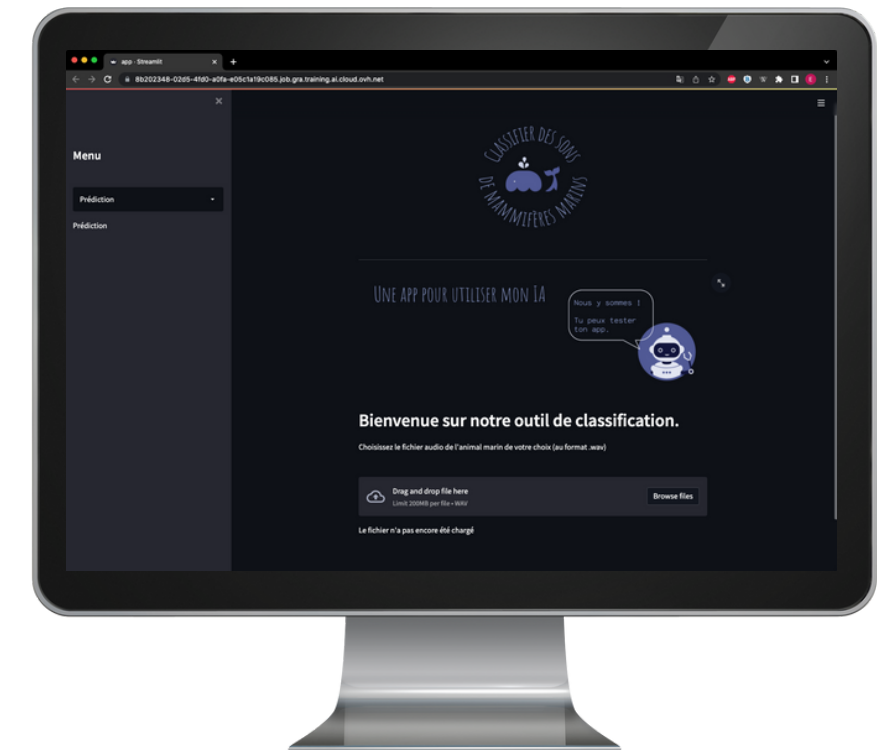
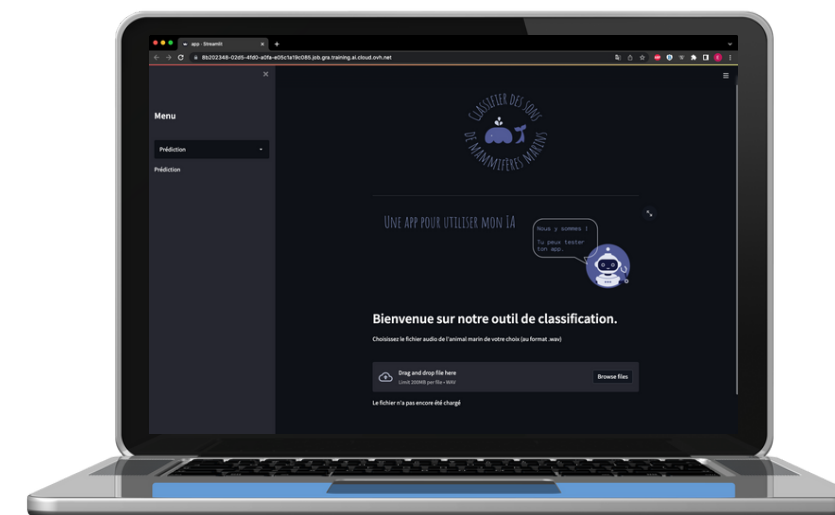
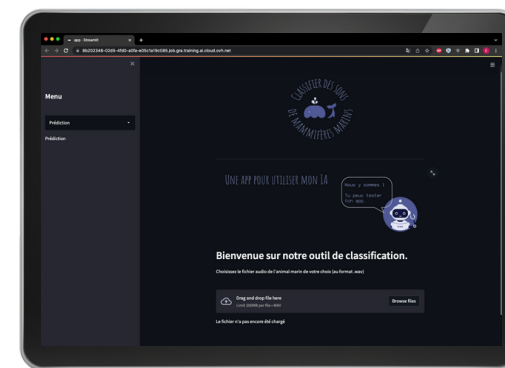


Tu veux ouvrir
ton notebook ?
Clic sur l'écran



UNE APP POUR UTILISER MON IA

Imaginez que votre IA soit accessible et utilisable par tout le monde...



COMMENT CRÉER UNE APP ?



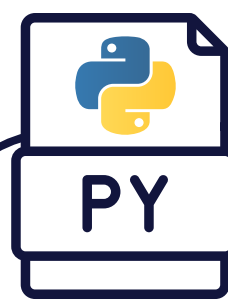
MODÈLE D'IA

- entraînement

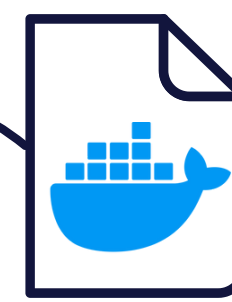
- exportation



Object Storage



CODE PYTHON

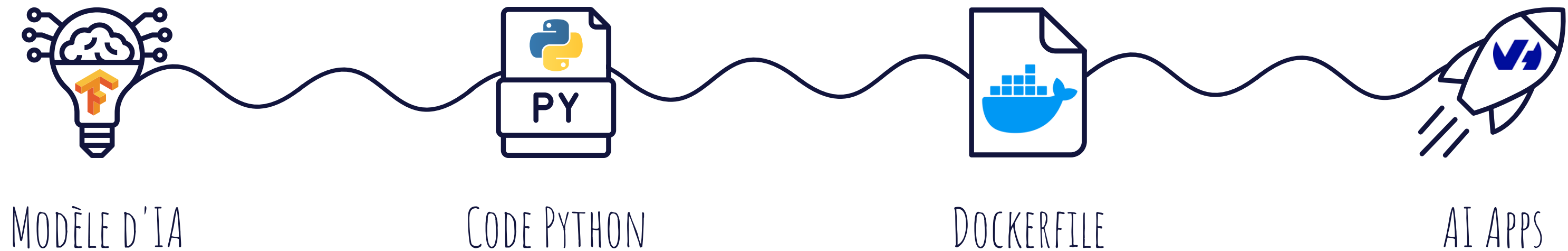


DOCKERFILE



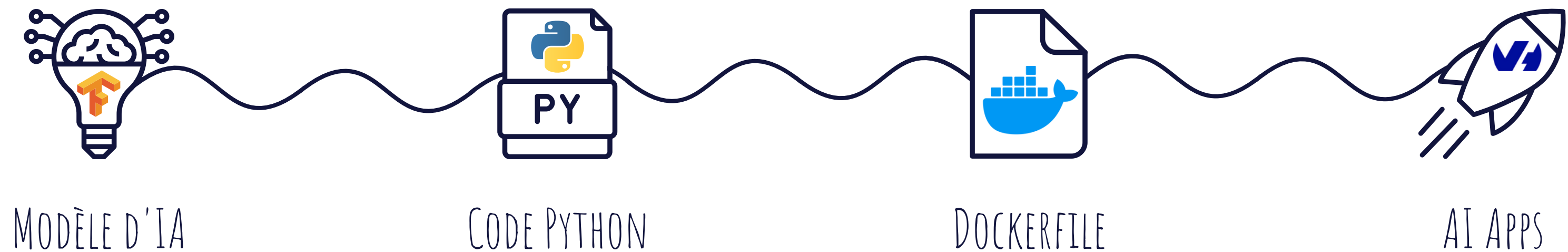
AI APPS

COMMENT CRÉER UNE APP ?



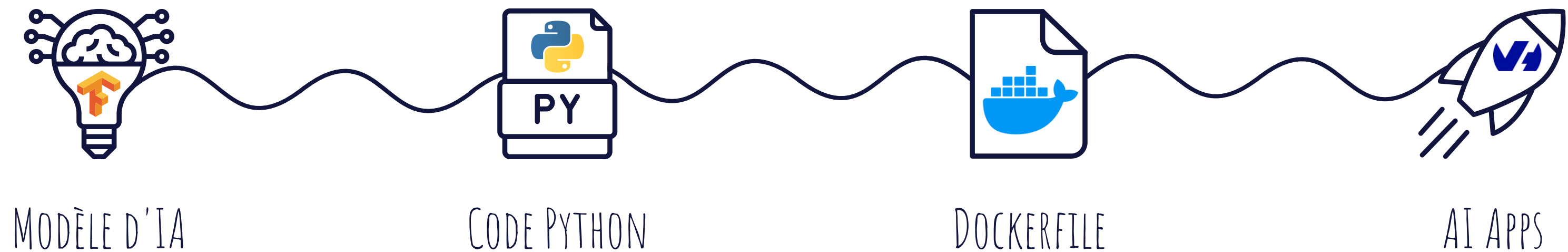
- 🏰 framework Streamlit
- 🐍 transformation du son
- 🐍 classification

COMMENT CRÉER UNE APP ?



- image Docker de base
- répertoire courant
- ↩ dépendances Python
- commande à exécuter

COMMENT CRÉER UNE APP ?

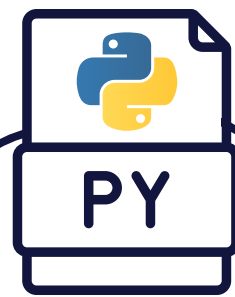


- image Docker personnalisée
- port d'accès du job
- CPU / GPU
- conteneur objet du modèle

COMMENT CRÉER UNE APP ?



MODÈLE D'IA



CODE PYTHON



DOCKERFILE

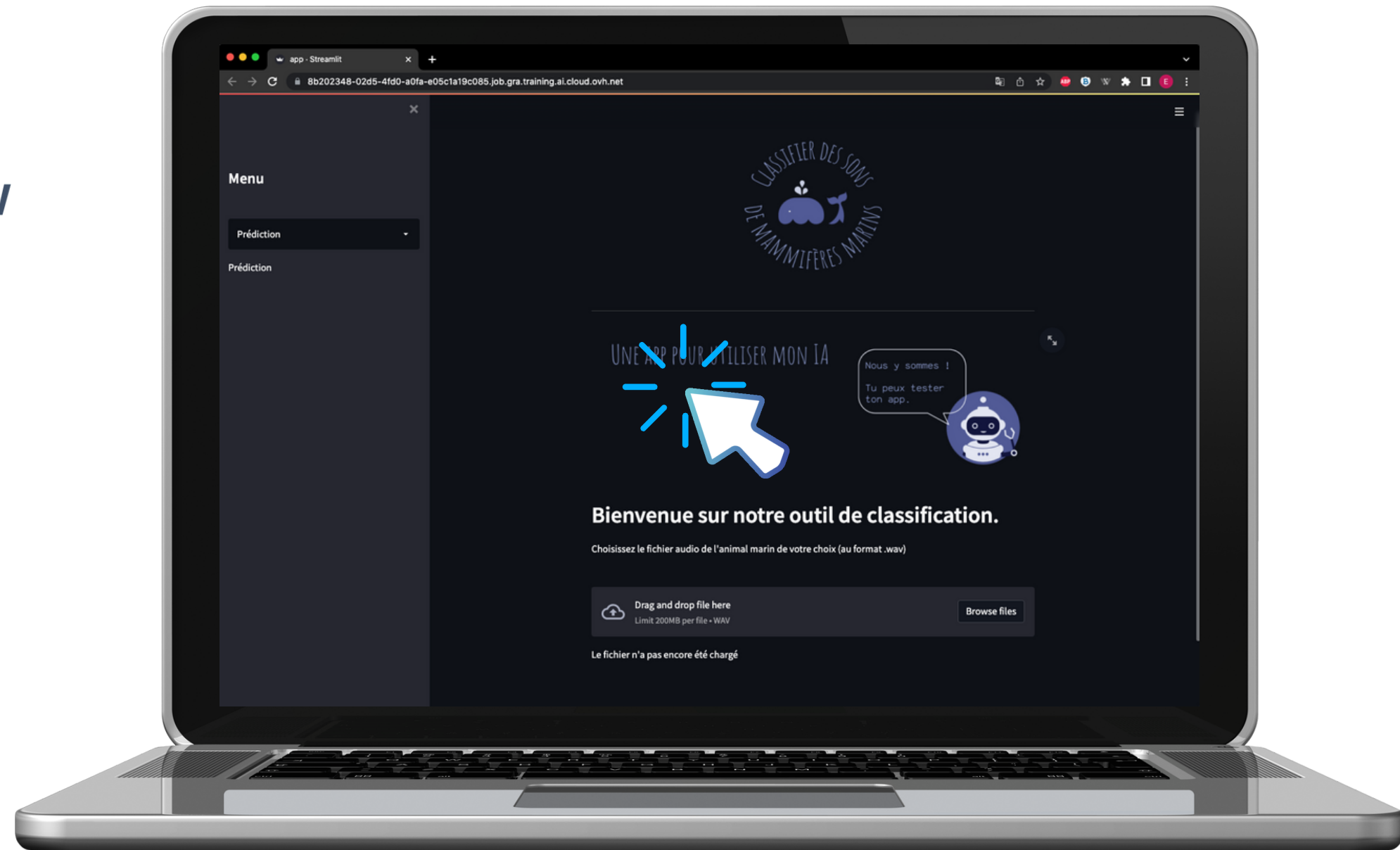
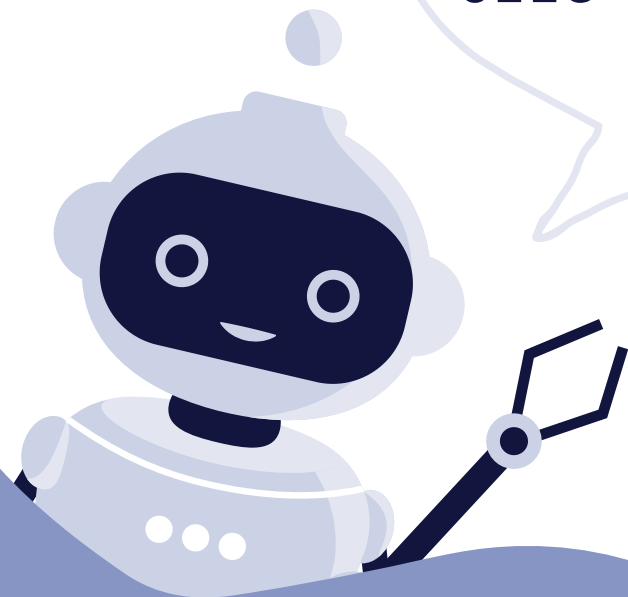


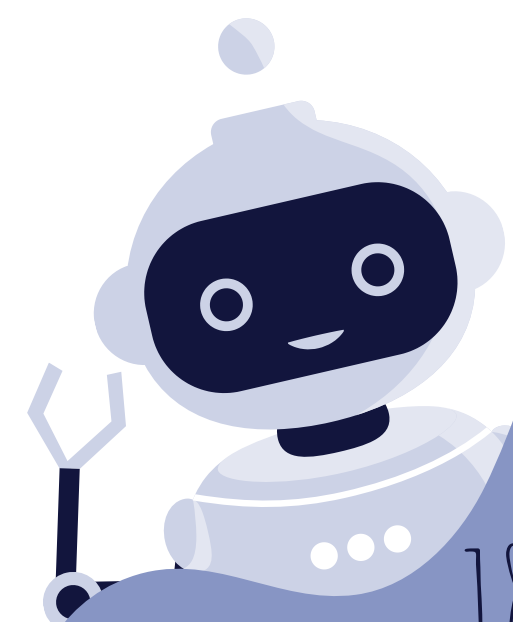
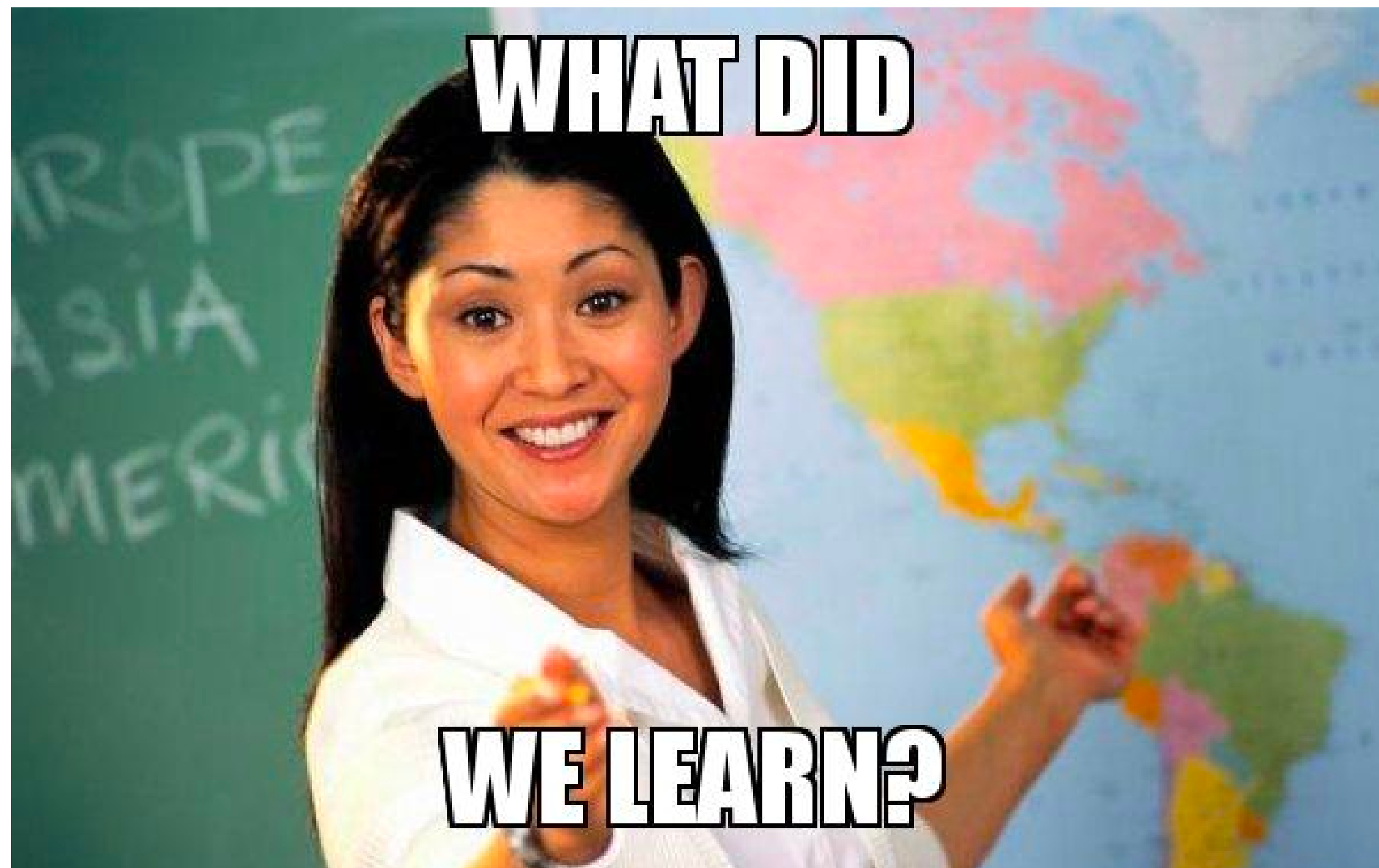
AI APPS

TEST DE MON APP



Tu veux tester
ton app ?
Clic sur l'écran





QU'AVONS NOUS APPRIS AUJOURD'HUI ?

NOTION D'IA

Datascience, Machine Learning, Deep Learning

PLACE DU SON EN IA

Environnement, musique, NLP

PRÉTRAITEMENT DES DONNÉES

Répartition, transformation, labélisation

AI NOTEBOOK

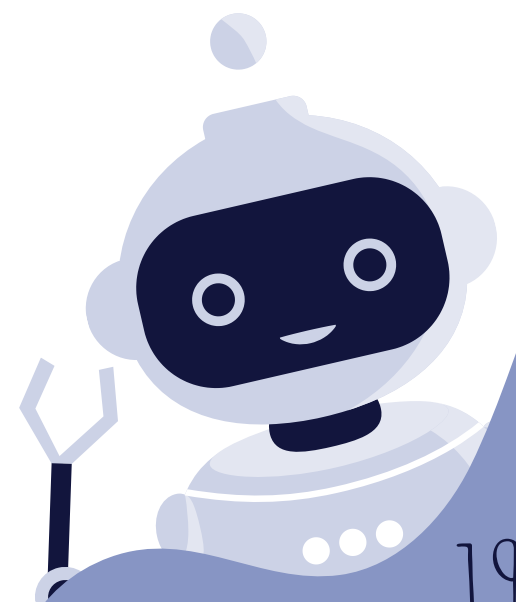
Développement de mon code étape par étape

ENTRAÎNEMENT D'UNE IA

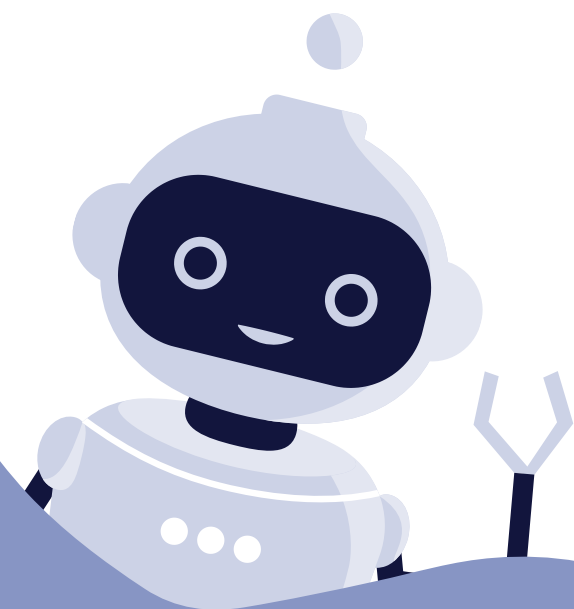
Réseau de neurones, exportation

APP POUR IA

Object Storage, Streamlit, Docker, AI Apps



MERCI !



POUR TESTER

 LE NOTEBOOK : <https://github.com/ovh/ai-training-examples/blob/main/notebooks/audio/audio-classification/notebook-marine-sound-classification.ipynb>

 LA DOC DU NOTEBOOK : <https://docs.ovh.com/gb/en/publiccloud/ai/notebooks/tuto-marine-mammal-sounds-classification/>

 L'APP : <https://github.com/ovh/ai-training-examples/tree/main/jobs/streamlit/audio-classification-app>

 LA DOC DE L'APP : <https://docs.ovh.com/gb/en/publiccloud/ai/training/tuto-streamlit-sounds-classification/>

 L'ARTICLE DE BLOG : <https://blog.ovhcloud.com/ai-notebooks-analyze-and-classify-sounds-with-ai/>

REPO GITHUB



RÉFÉRENCES

- **TensorFlow** : <https://www.tensorflow.org/guide?hl=fr>
- **JupyterLab** : <https://jupyter.org/>
- **Pandas** : <https://pandas.pydata.org/>
- **Sklearn** : <https://scikit-learn.org/stable/>
- **Numpy** : <https://numpy.org/>
- **Matplotlib** : <https://matplotlib.org/>
- **Dataset des sons de mammifères marins** : <https://www.kaggle.com/datasets/shreyj1729/best-of-watkins-marine-mammal-sound-database>
- **Informations sur la dataset** : <https://cis.whoi.edu/science/B/whalesounds/index.cfm>
- **Classification de musique avec un CNN** : <https://blog.clairvoyantsoft.com/music-genre-classification-using-cnn-ef9461553726>
- **Classification des genre musicaux** : <https://towardsdatascience.com/music-genre-classification-with-python-c714d032f0d8>
- **Pré-traitement des données en Machine Learning** : <https://towardsdatascience.com/introduction-to-data-preprocessing-in-machine-learning-a9fa83a5dc9d>