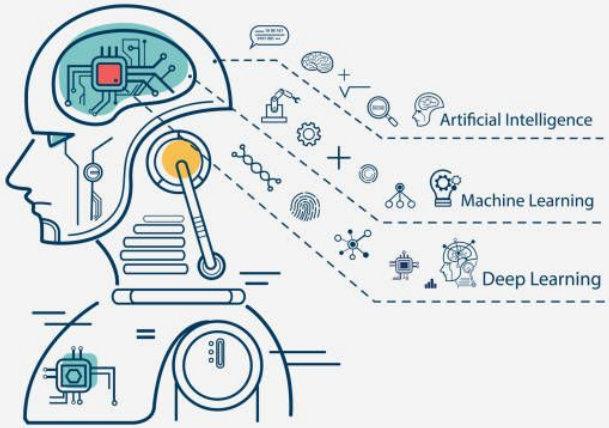
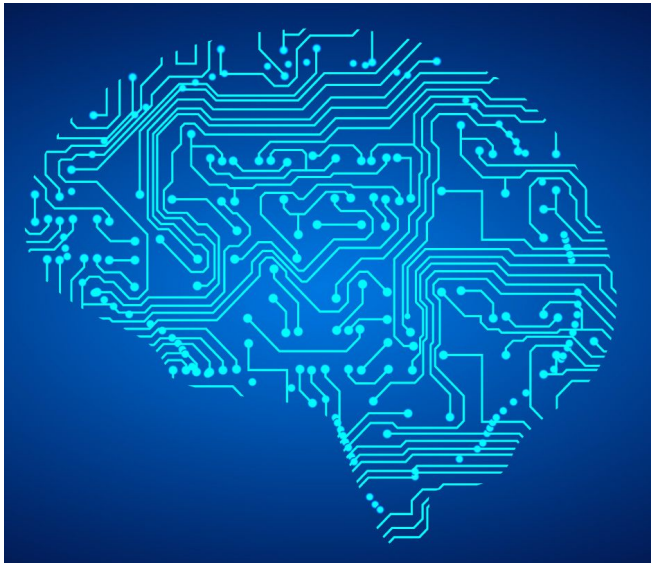




Machine Learning Overview

lesson 1

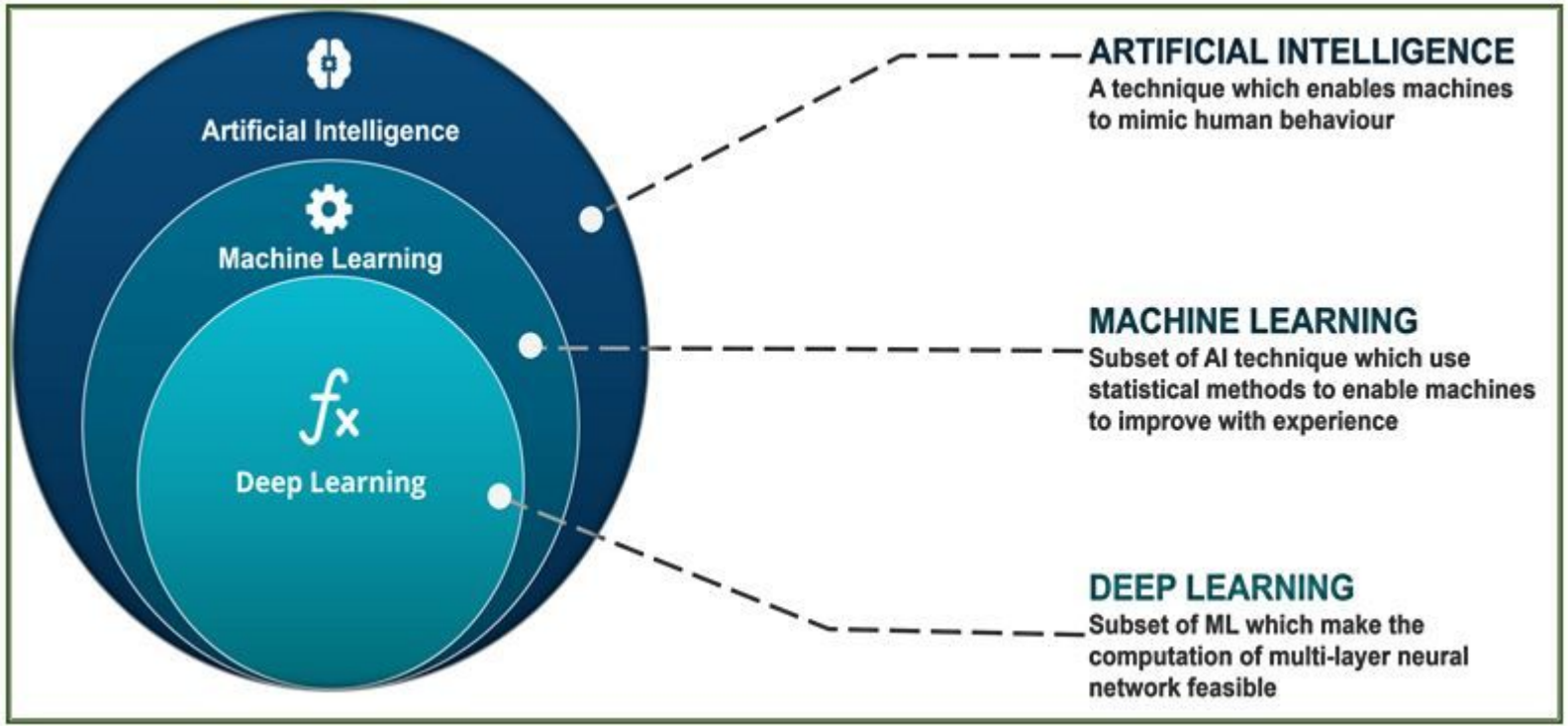
Mously DIAW

ENSAE - 16/12/2021

Summary

- What is Artificial Intelligence (AI) ?
- What is Machine Learning (ML) ?
- Deep Learning (DL) overview
- How to learn Data Science ?
- Project description
- Questions

Artificial intelligence domain: key concepts ?



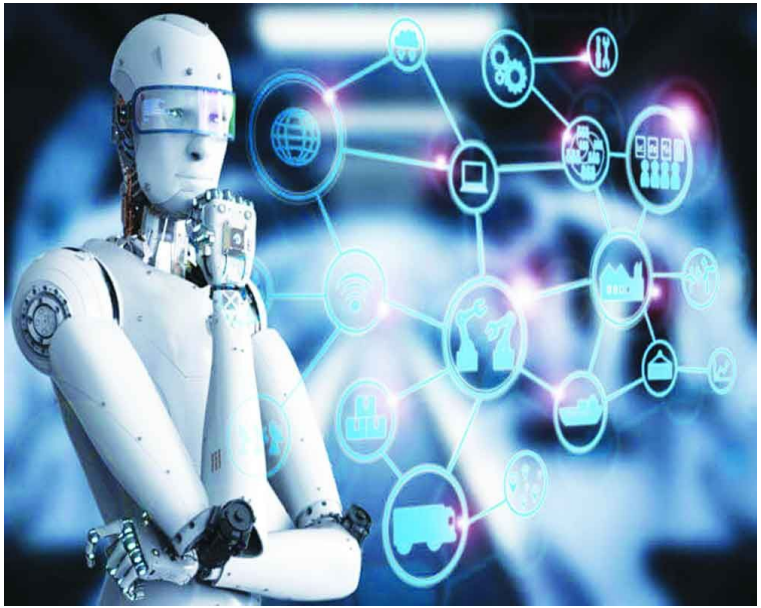
Artificial Intelligence (AI)

AI, is all the theories and techniques implemented in order to create machines capable of simulating human intelligence

The term "artificial intelligence", was created by John McCarthy (1950s)

According to Marvin Lee Minsky, AI is:

“the science of making machines do things that would require intelligence if done by men”



What is Machine Learning ?

Introduction to Machine Learning (ML)

x	y
1	1
2	2
3	3
4	4
5	5
6	6

What do you think of mathematical relation exist between x & y ?

y = x

Introduction to Machine Learning

x	y
1	10
2	20
3	30
4	40
5	50
6	60

What do you think of mathematical relation exist between x & y ?

y = 10 * x

Introduction to Machine Learning

x1	x2	y
1	2	3
2	3	5
3	4	7
4	5	9
5	6	11
6	7	13

What do you think of mathematical relation exist between x & y ?

$$y = x1 + x2$$

Introduction to Machine Learning

x1	x2	x3	x4	...	x1000	y
1	2					3
2	3					5
3	4					7
4	5					9
5	6					11
6	7					13

... Finding the relationship between X (independent variables) et y (dependent variable)

y = f(x1, x2, x3, ..., x1000)

Introduction to Machine Learning

$$x1, x2, x3, \ldots, x1000$$

$$\begin{pmatrix} 0 & 3 & 1 & 0 & 2 & 3 & 8 & 1 & 1 & 3 \\ 1 & 1 & 0 & 0 & 7 & 1 & 2 & 2 & 3 & 3 \\ 1 & 2 & 2 & 0 & 0 & 6 & 7 & 1 & 2 & 2 \\ 1 & 2 & 3 & 10 & 0 & 4 & 6 & 1 & 0 & 5 \\ 3 & 2 & 2 & 1 & 4 & 3 & 2 & 1 & 6 & 0 \\ 7 & 4 & 4 & 5 & 3 & 9 & 6 & 1 & 6 & 1 \\ 7 & 1 & 1 & 5 & 2 & 8 & 9 & 1 & 3 & 6 \\ 5 & 0 & 1 & 6 & 2 & 0 & 0 & 0 & 1 & 5 \\ 1 & 6 & 3 & 3 & 4 & 6 & 2 & 0 & 1 & 1 \\ 1 & 2 & 2 & 4 & 1 & 1 & 3 & 0 & 8 & 2 \end{pmatrix}$$

y = cat

Machine Learning formally

In 1959, Arthur Samuel define ML as:

> **“field of study that gives computers the ability to learn without being explicitly programmed”**

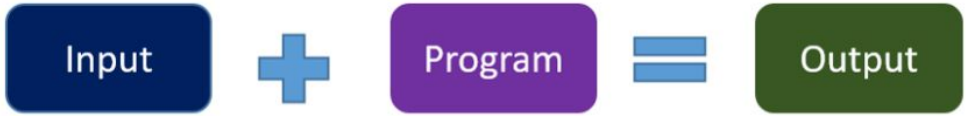
- Machine learning is finding such complex mathematical relation between X & Y
- Machine learning helps to find hidden insight about data without doing explicit programming
- ML can be understood as computational methods that use experience to improve performance or to make accurate predictions.

Example of ML Use cases

- Object Recognition
- Fraud Detection
- Product Recommendations
- Digital Marketing
- Price prediction
- Face Detection
- Text Recognition
- Translate Text
- Email filtering

Machine Learning vs Traditional system of Computing

Traditional programming



is a manual process - meaning a person (programmer) creates the program with business/solution rules

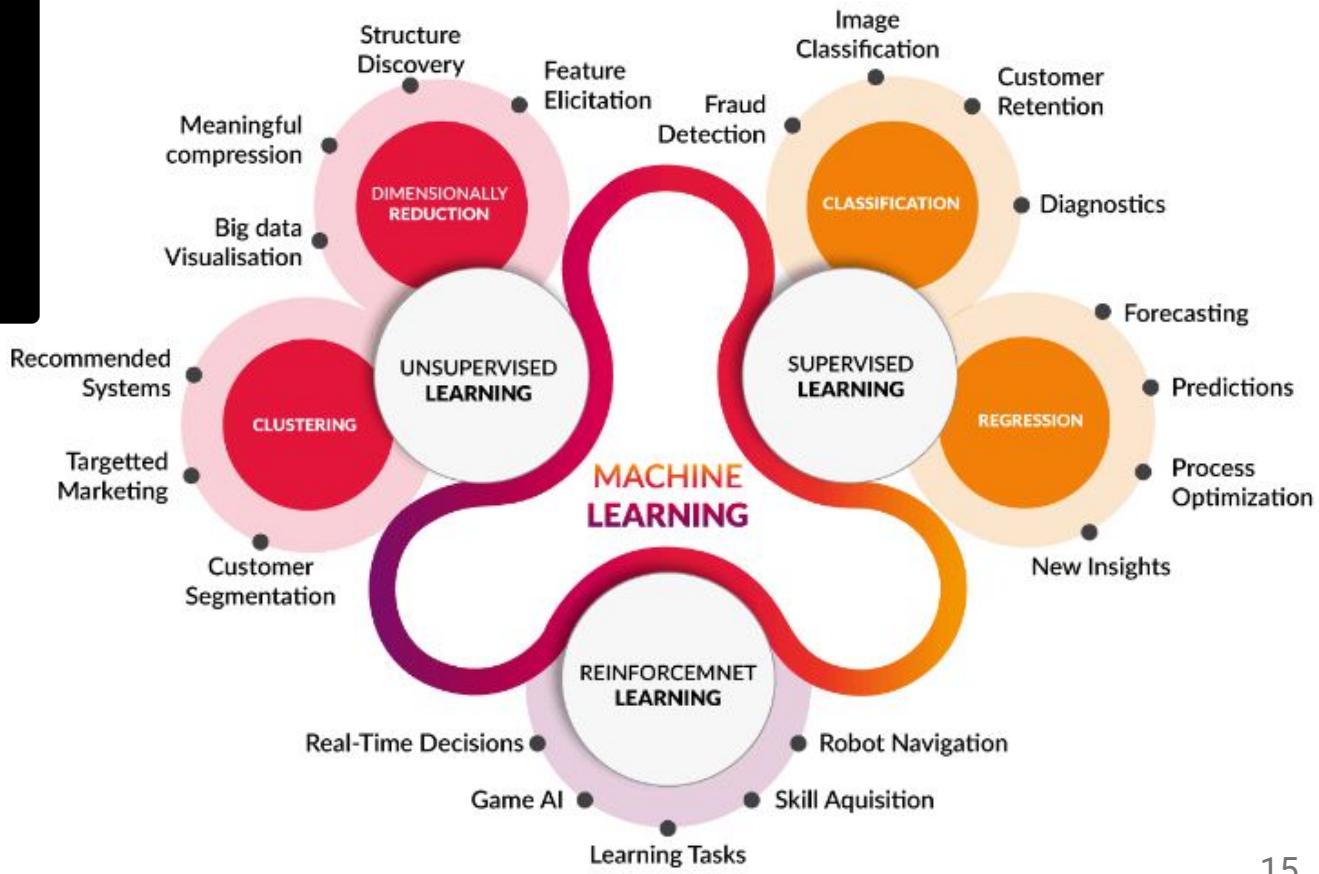
---

Machine Learning

the input data and output are fed to an algorithm to create a program.

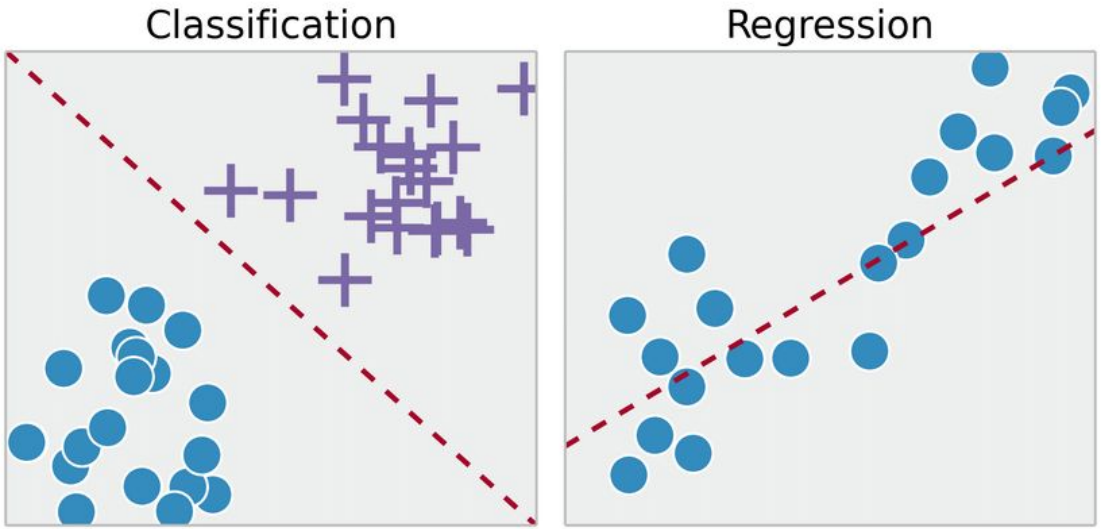
Types of Machine Learning

Types of Machine Learning



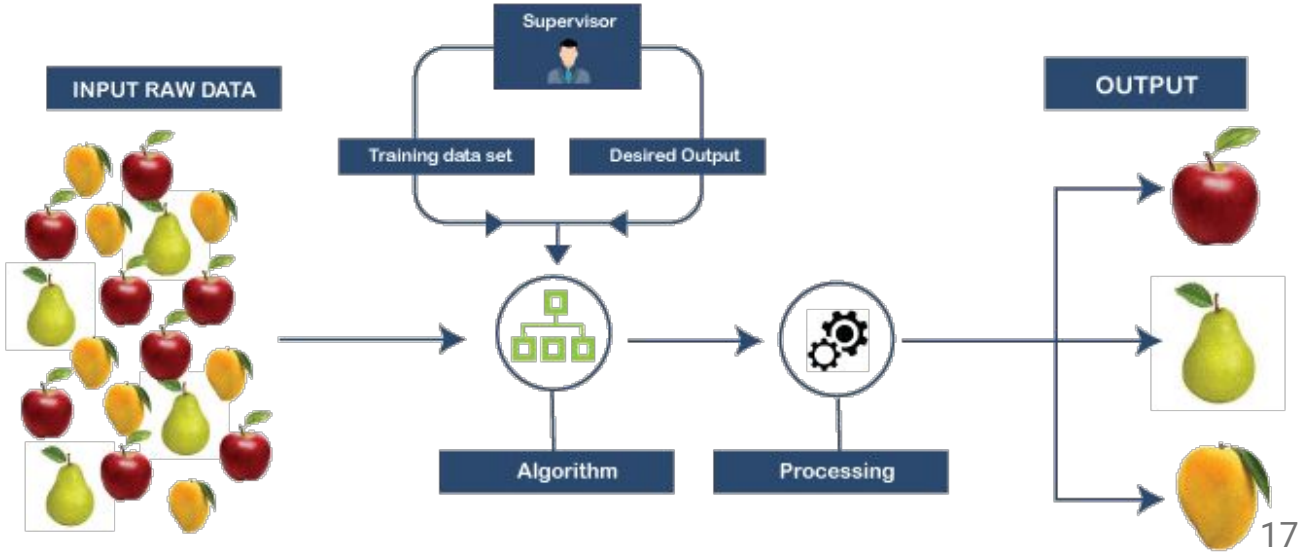
Supervised Learning

- task of learning a function that maps an input to an output based on example.
- Models are trained using labelled dataset (i.e., each input is paired with some output)
- It is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately
- All d**ata is labeled** and the algorithms learn to predict the output from the input data.



Supervised Learning

- **Classification** uses an algorithm to accurately assign test data into specific categories.
    - Customer sentiment analysis, spam detection, image detection
- **Regression** is used to understand the relationship between dependent (numeric) and independent variables.
    - Dynamic price: (hostel, booking,

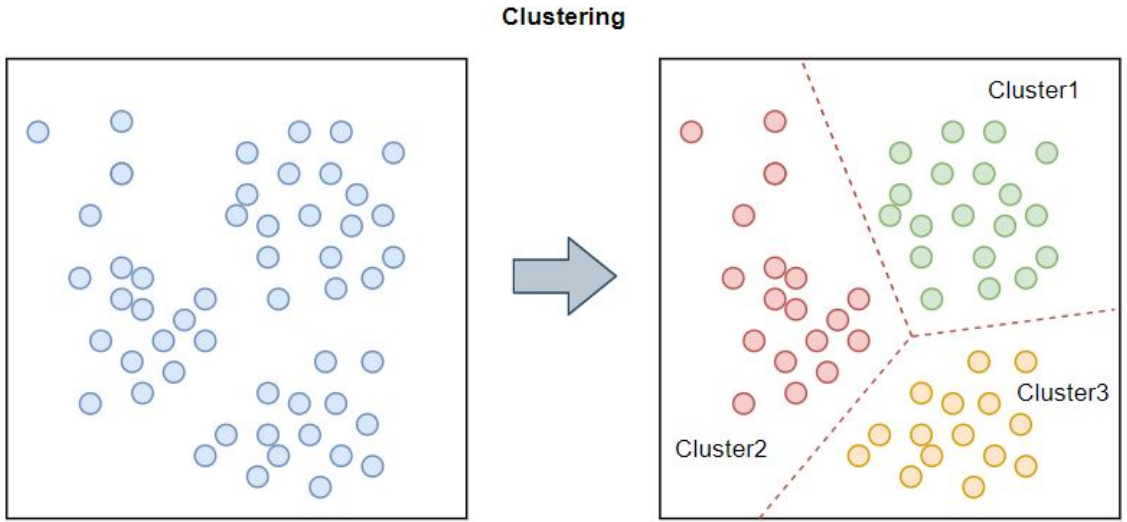


Unsupervised Learning

- Machine learning that are designed to learn by examples, i.e., It maps the input to an output based on previous input-output pairs. It is trained with labelled data
- Models are trained using labelled dataset (i.e., each input is paired with some output)
- All **data is unlabeled** and the algorithms learn to inherent structure from the input data.

Unsupervised Learning

- **Clustering** is a data mining technique for grouping unlabeled data based on their similarities or differences
- **Association** is another type of unsupervised learning method that uses different rules to find relationships between variables in a given dataset
- **Dimensionality** reduction is a learning technique used when the number of features (or dimensions) in a given dataset is too high



Unsupervised Learning

Network graphs

01

Explanations

network models or forecasting real-world behavior

multiple people connected together

risky pattern detection

help to see strange patterns

02

Examples

fraud detection

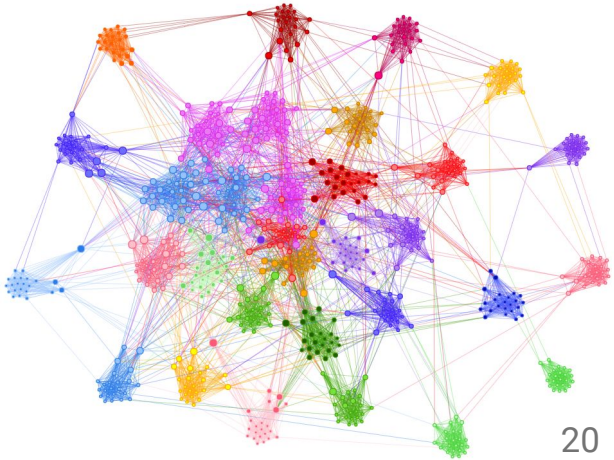
recommendation

social network analysis

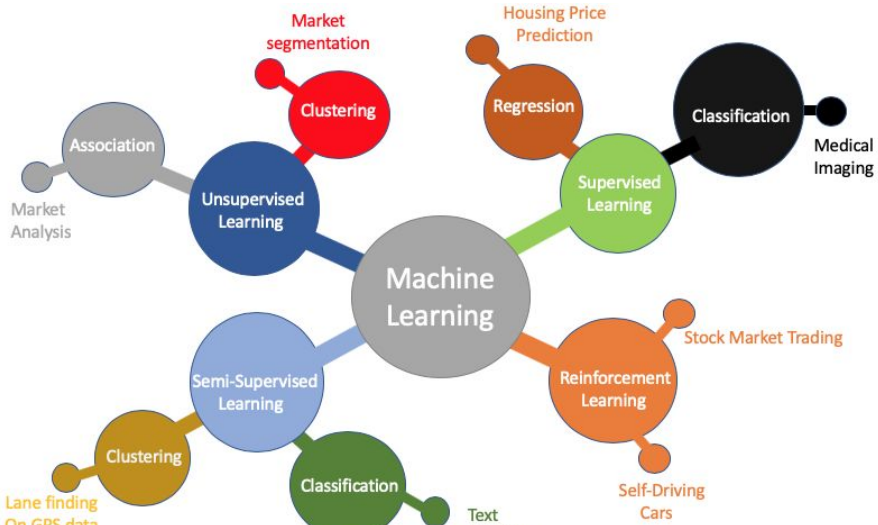
03

Python Packages

networkx
neo4j
pyvis
visdcc
plotly



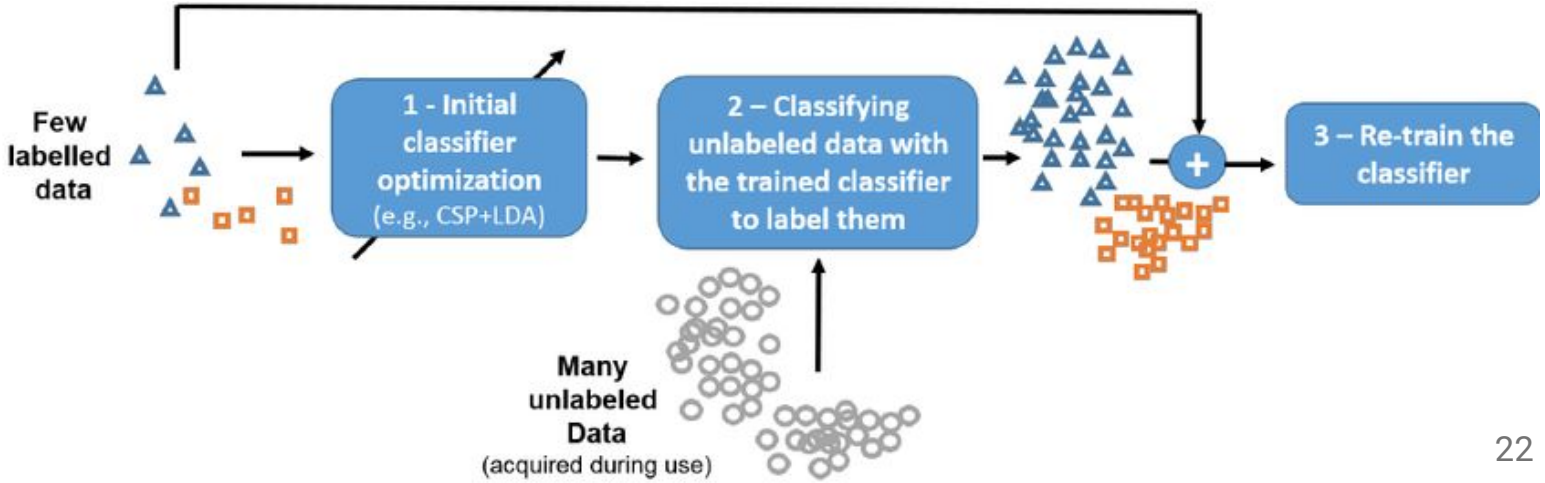
Semi supervised Learning



- **semi-supervised learning:** occurs when only part of the given input data has been labeled.
- you have a large amount of input data (X) and only some of the data is labeled (Y) are called semi-supervised learning problems.
    - A good example is a photo archive where only some of the images are labeled, (e.g. dog, cat, person) and the majority are unlabeled.
- Some data is labeled but most of it is unlabeled and a mixture of supervised and unsupervised techniques can be used.

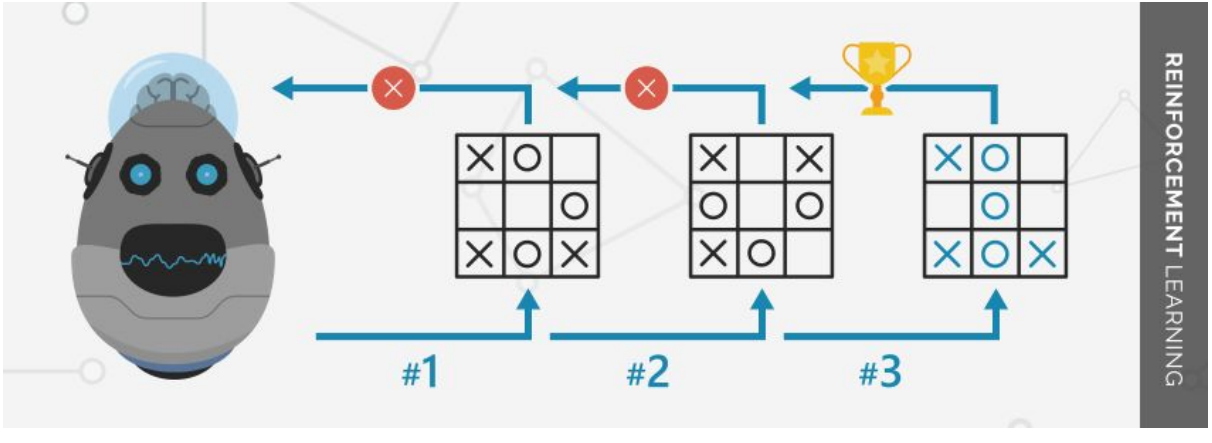
Semi supervised Learning

- **semi-supervised learning:** occurs when only part of the given input data has been labeled.
- it lies between Supervised and unsupervised learning
- you have a large amount of input data (X) and only some of the data is labeled (Y) are called semi-supervised learning problems.
    - A good example is a photo archive where only some of the images are labeled, (e.g. dog, cat, person) and the majority are unlabeled.
- Some data is labeled but most of it is unlabeled



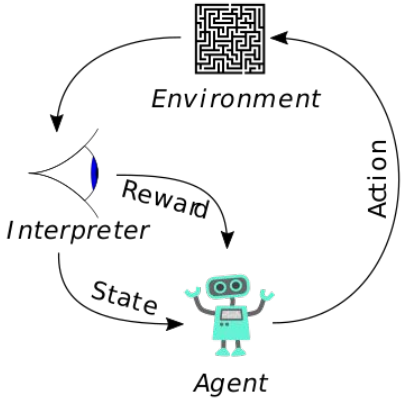
Reinforcement Learning

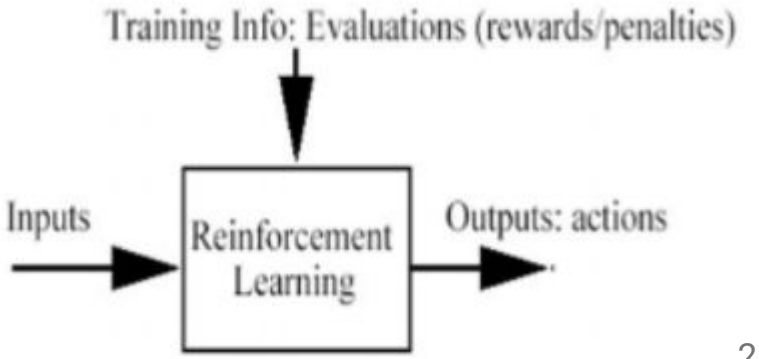
- Focus on making decisions based on previous experience
- The purpose of RL is for the agent to learn an optimal, or nearly-optimal, policy that maximizes the "reward function"
- Policy-making with feedback, there are no training data
    - An example of RL is a playing game
- Markov decision process (MDP) in order to maximizes the expected cumulative reward.
- the focus is on finding a balance between exploration (of uncharted territory) and exploitation



Reinforcement Learning

- RL is often used to make strategies
- is like unsupervised ML in that the training data is also unlabelled
- The purpose of RL is for the agent to learn an optimal, or nearly-optimal, policy that maximizes the "reward function"
- Policy-making with feedback, there are no training data
    - An example of RL is a playing game
- Markov decision process (MDP) in order to maximizes the expected cumulative reward.
- the focus is on finding a balance between exploration (of uncharted territory) and exploitation





Type of ML - Summary

Supervised Learning	Unsupervised Learning	Semi-supervised Learning	Reinforcement Learning
Data has knows labels/outputs Example Algorithms: • Linear Regression: • Support Vector Machines (SVM) • Decision Tree • Stochastic Gradient Descent • Ensemble methods • Nearest Neighbors • Neural network models	Unlabeled data Focus on finding patterns and gaining insight from the data Example Algorithms: • K-means clustering • Gaussian mixture models • Vector quantization • Matrix factorization problems (PCA, ICA, LDA) • Novelty & Outlier detection	Build a model through a mix of labeled & unlabeled data It stands between Supervised & unsupervised learning Example Algorithms • Generative adversarial networks • Self training classifier	Focus on making decisions based on previous experience making decisions sequentially Example Algorihms: • Model free RL (Q-learning, policy optimization) • Model-based RL (learn/given) the model
Use cases: • sales forecasting • risk assessment • image classification • Insurance underwriting • Fraud detection	Use cases: • Customer segmentation • Word associations • Searcher intent • Recommender systems • predictive maintenance	Use cases: • Medical predictions • Audio & video manipulation • Natural language preprocessing	Use cases: • Game AI • Self driving cars • text summarization • question answering • robotics manipulation

Deep learning overview

Deep Learning overview

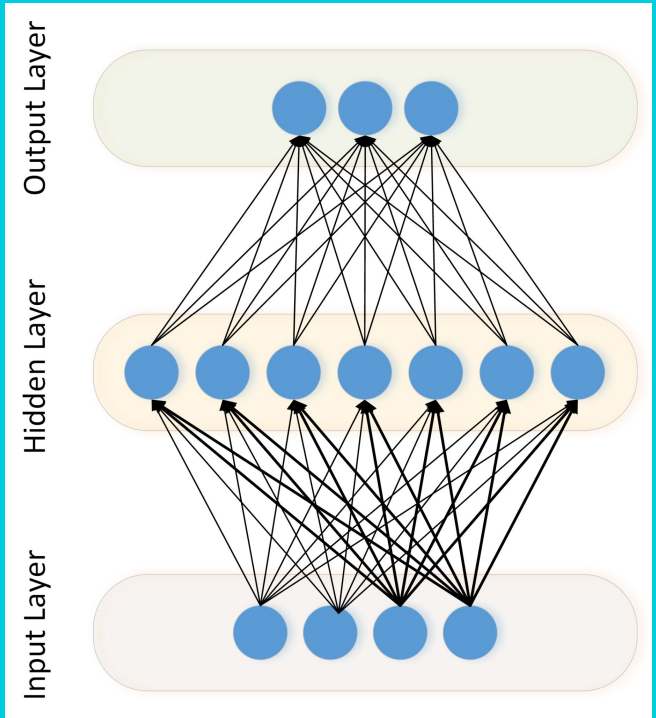
Inspired by our brain's network of neuron

DL was first theorized in the 1980s

A deep neural network (DNN) is an artificial neural network (ANN) with multiple layers between the input and output layers

Such systems learn (progressively improve their ability) to do tasks by considering examples

There are different types of neural networks but they always consist of the same components: neurons, synapses, weights, biases, and functions



Deep Learning overview

These components functioning similar to the human brains and can be trained like any other ML algorithm.

Deep learning requires large amounts of **data**

Deep learning requires substantial **computing power**. High-performance GPUs have a parallel architecture that is efficient for deep learning

Most common ways people use DL:

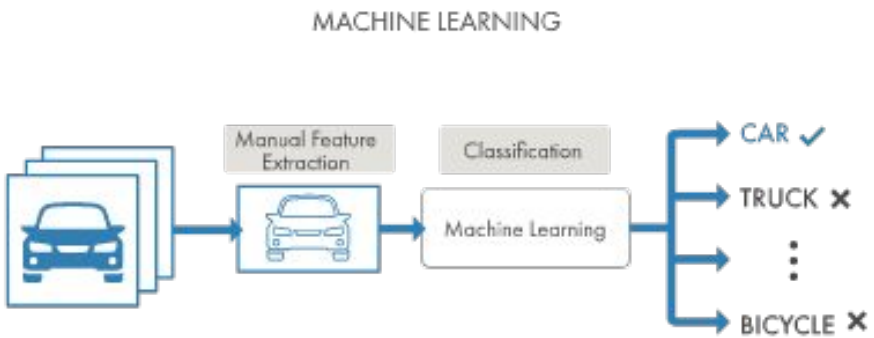
- Training a model from scratch
- Transfert learning
- Feature extraction

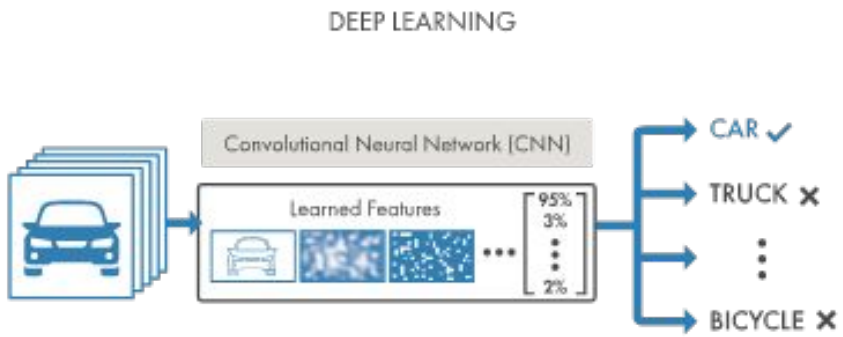
Examples of use cases:

- Automated Driving
- Image detection
- Medical Research
- Machine translation

Machine Learning (ML) vs Deep Learning (DL)

- Deep learning is a specialized form of machine learning
- ML workflow starts with relevant features being manually extracted from images
- With DL, feature extraction and modeling steps are automatic
- Another key difference is deep learning algorithms scale with data (more data, better performance), whereas shallow learning converges

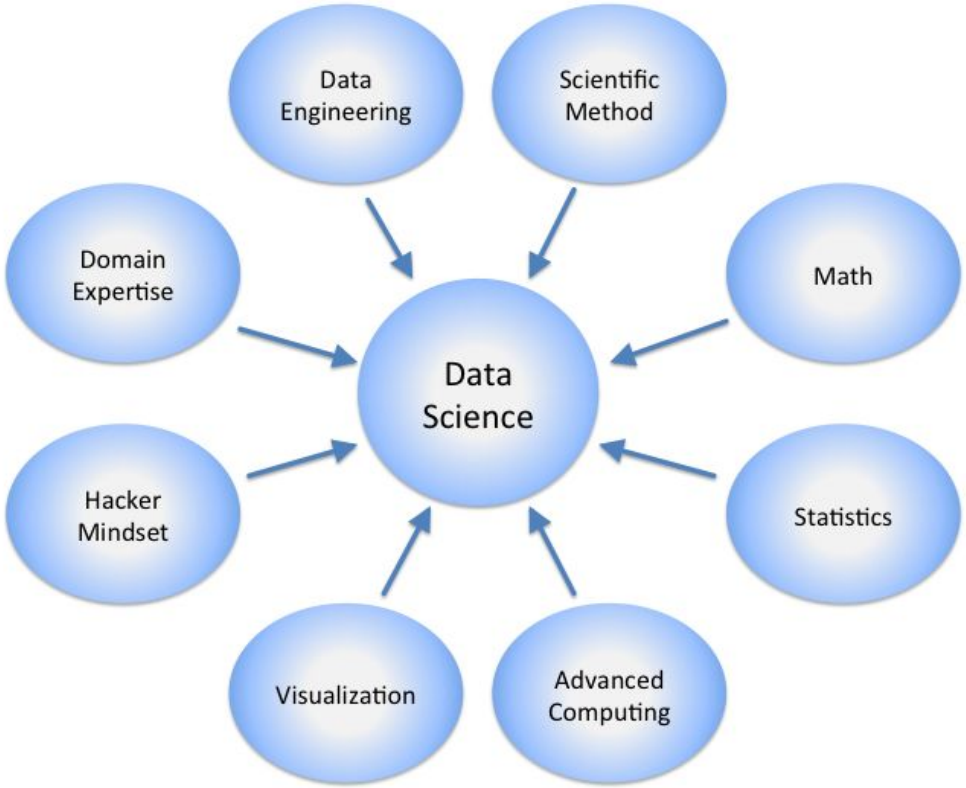




How to learn Data Science ?

Learn to Data Science

Get good at stats, maths, machine learning & programming



Learn to Data Science

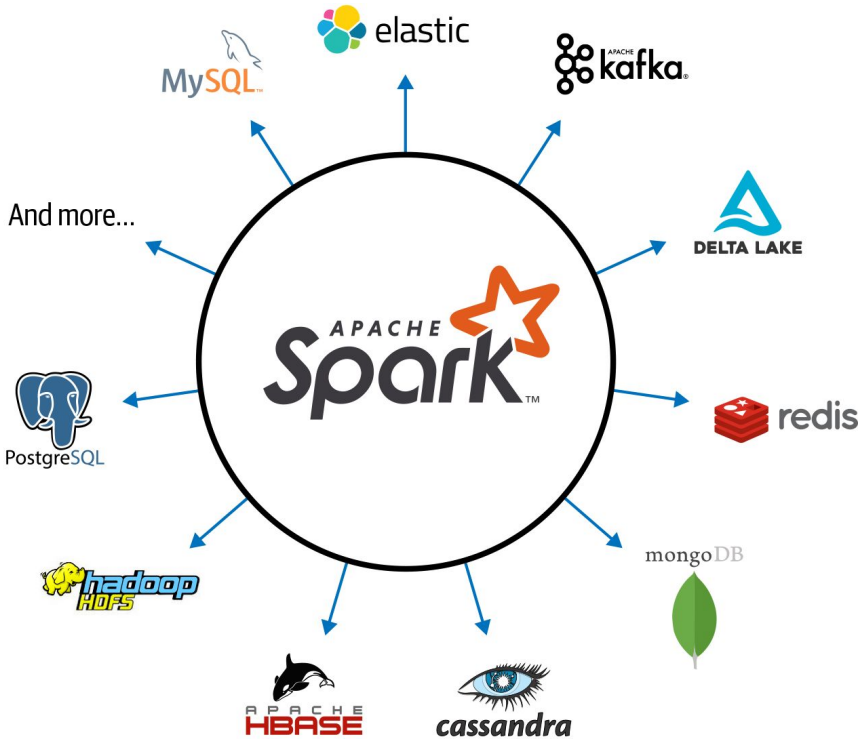
Learn to code

Learn to Data Science

Understand database

Learn to Data Science

Level Up with Big data



Learn to Data Science

Grow, connect & learn

kaggle

meetup

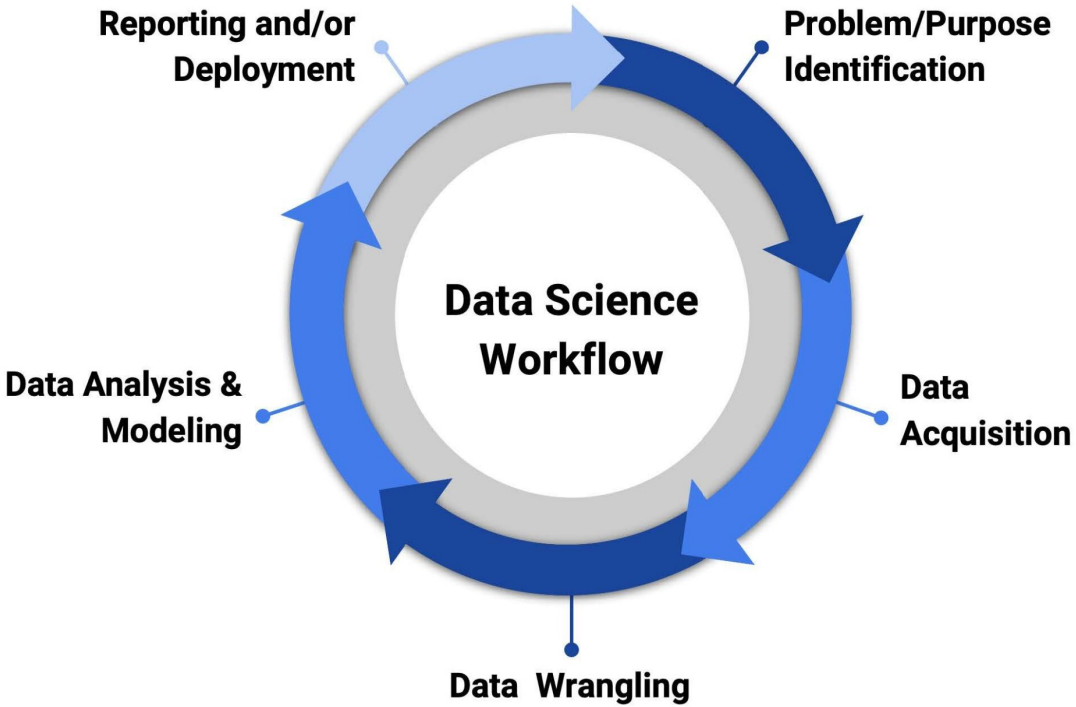
moodle

coursera

stackoverflow

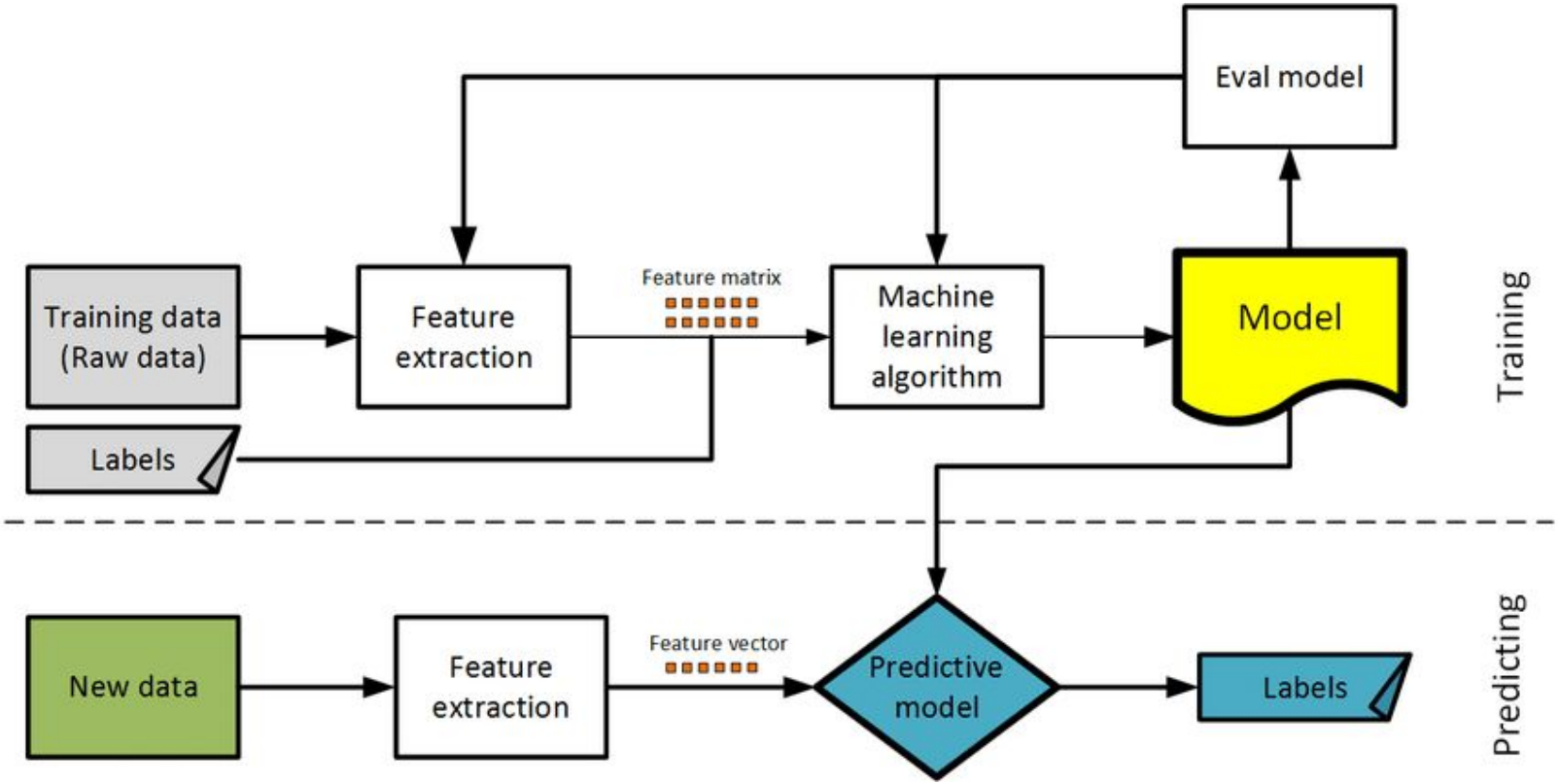
Learn to Data Science

Explore Data Science workflow (1/2)



Learn to Data Science

Explore Data Science workflow (2/2)



Project description

Project outputs

Présentation: 20 min

Discussion: 5 à 10 min

Github

Pour faciliter votre passage, déposez sur github, dans un dossier nommé “NomDuProjet_NumeroGroup”, tous les livrables du projet. Chaque livrable doit être nommé selon les actions effectuées, par exemple “notebookanalyse”, “notebookessais”, et ainsi de suite.

Support de présentation

- Rappel de la problématique
- Présentation du jeu de données
- Présentation de l’analyse pré-exploratoire du jeu de données et vos conclusions sur la pertinence de l’usage du jeu de données
- Présentation des différentes pistes de modélisation effectuées et du modèle final sélectionné

THANK YOU

Questions ?

References

https://www.quantmetry.com/blog/une-petite-histoire-du-machine-learning/

https://www.vxchnge.com/blog/machine-learning-use-cases

https://developers.google.com/learn/topics/on-device-ml?utm_source=google&utm_medium=cpc&utm_campaign=ODML-TensorFlow&utm_content=AI&gclid=Cj0KCQiAnuGNBhCPARIsACbnLzr3b-x3mVRirkkEy1Q4JkxoncQz5rO1Eq31lOJpGE652glDYa-j330aArFaEALw_wcB

https://analyticsindiamag.com/how-do-machine-learning-algorithms-differ-from-traditional-algorithms/

https://algorithmia.com/blog/machine-learning-use-cases

https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/

https://www.ibm.com/cloud/learn/supervised-learning

https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning

https://algorithmia.com/blog/semi-supervised-learning

https://en.wikipedia.org/wiki/Reinforcement_learning

https://theconversation.com/what-is-machine-learning-76759

https://towardsdatascience.com/getting-started-with-graph-analysis-in-python-with-pandas-and-networkx-5e2d2f82f18e