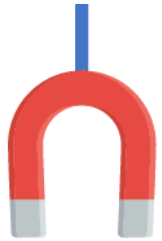
AI FROM ZERO TO HERO

Taghouti Ghofrane

WEB SCRAPING

scrapinghub



WEB SCRAPING AVEC PYTHON

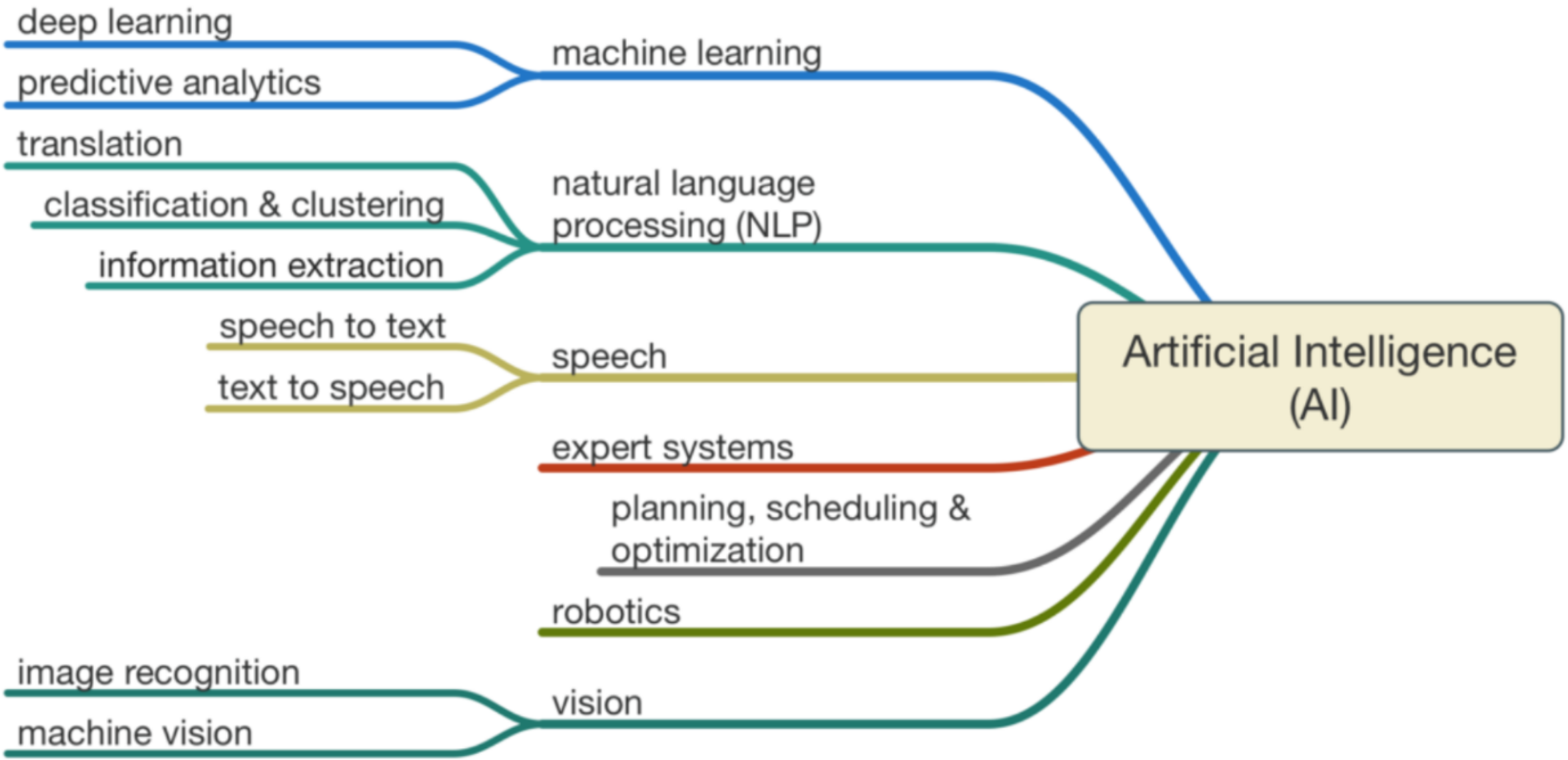
- **le développeur du scraper analyse le code source HTML** de la page à laquelle il s'intéresse.
- le code contient des schémas clairs permettant d'extraire les informations désirées.

Le scraper est programmé pour ces schémas. Le reste du travail est effectué **par le scraper de façon automatisée** et consiste à :

1. consulter un site Internet à une adresse URL,
2. extraire automatiquement les données structurées selon les schémas,
3. rassembler, enregistrer, analyser, combiner les informations extraites, etc.

LES BESOINS DE WEB SCRAPING AVEC PYTHON

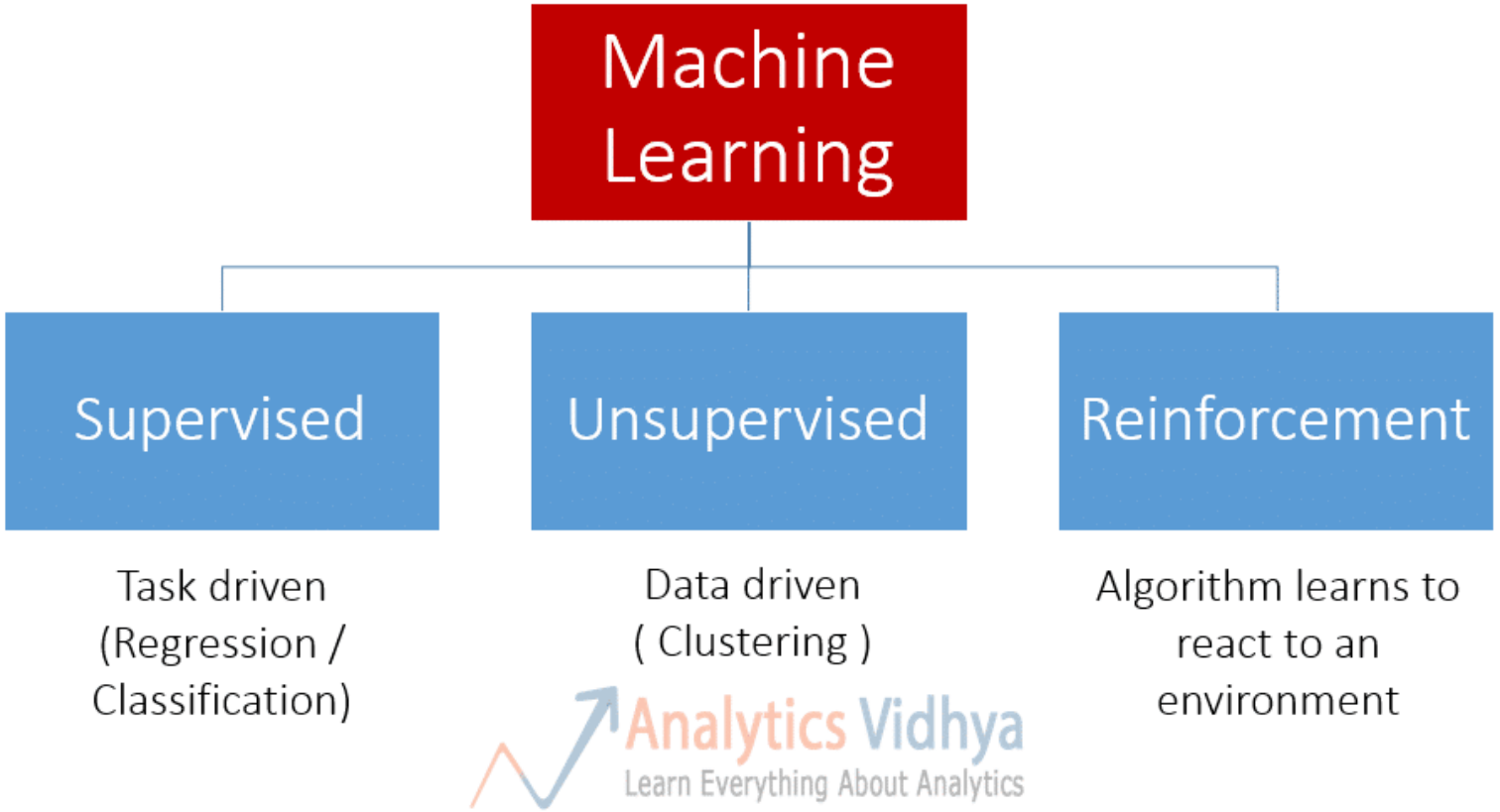
- Python (Anaconda / VS code)
- Selenium (Package Python)
- Lien = https://chromedriver.chromium.org/downloads
- Lien = https://chromedriver.chromium.org/getting-started


Artificial Intelligence (AI)
machine learning
deep learning
predictive analytics
natural language processing (NLP)
translation
classification & clustering
information extraction
speech
speech to text
text to speech
expert systems
planning, scheduling & optimization
robotics
vision
image recognition
machine vision

M L :

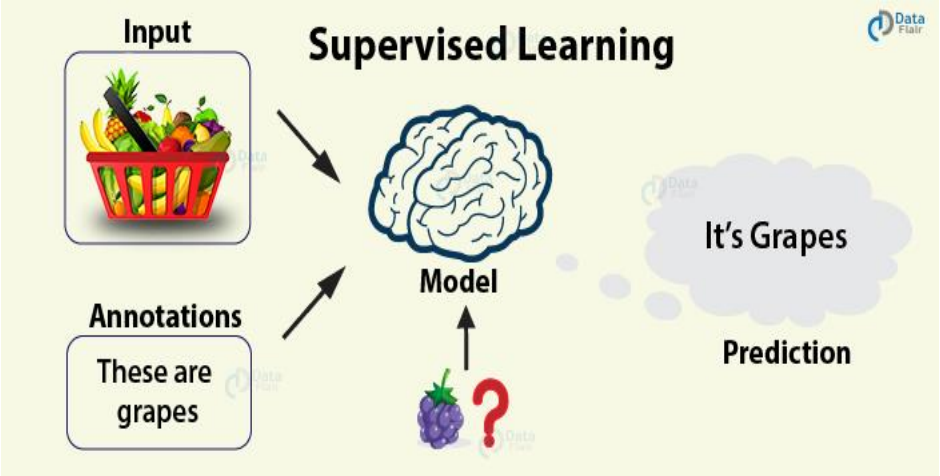
Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

Types of Machine Learning



SUPERVISED MACHINE LEARNING

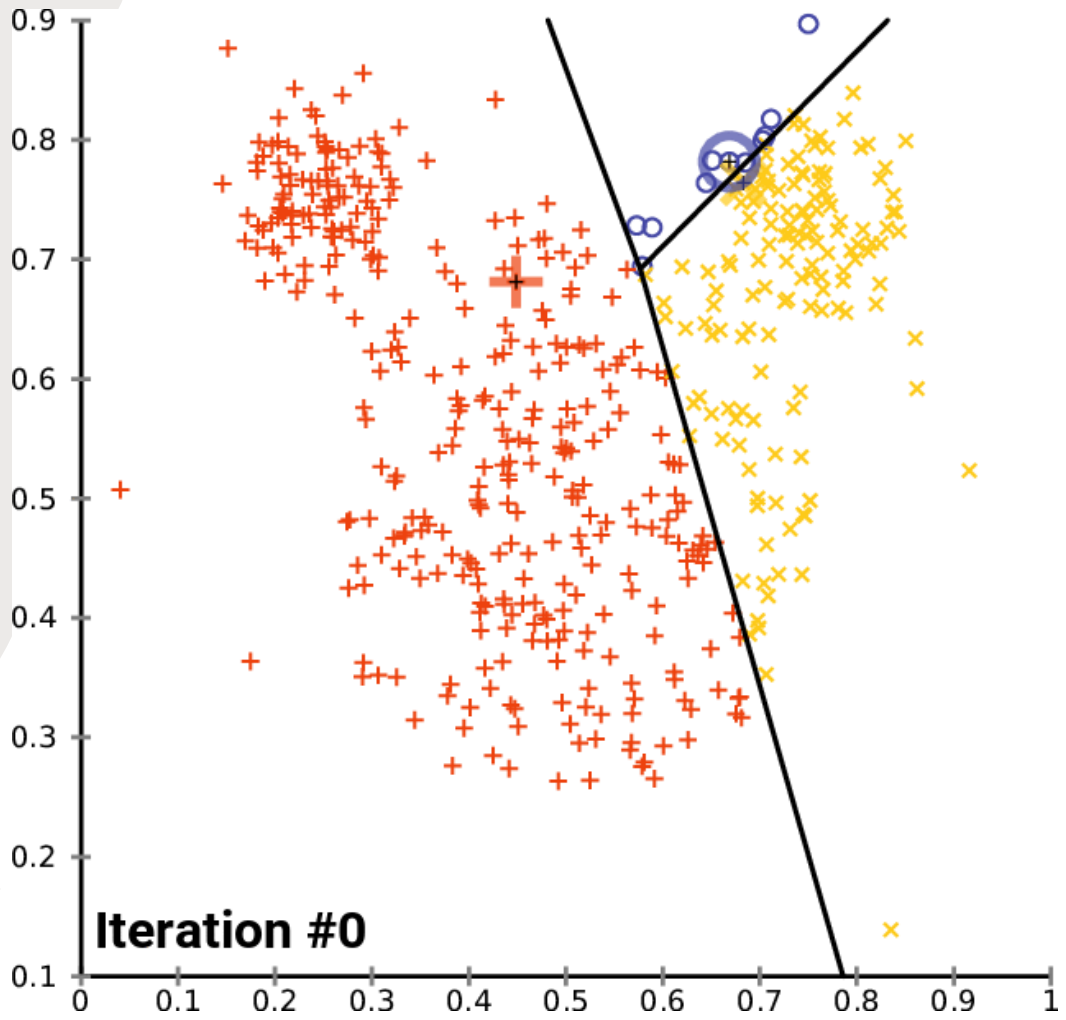
- **supervised** by someone.
- the machine uses **data** which is **already tagged** with the correct answer.
- The algorithm analyzes the **training data** and produces a **correct outcome** from **labeled data**.
- **regression**



UNSUPERVISED LEARNING

- **sort of algorithm** that learns patterns from **untagged data**.
- The machine learning algorithm acts on **information** without **guidance**.

- **Example :Customer Segmentation with Machine Learning (clustering not classification)**



SEMI_SUPERVISED LEARNING

Supervised learning aims to learn a function that, given a sample of data and desired outputs, approximates a function that maps inputs to outputs.

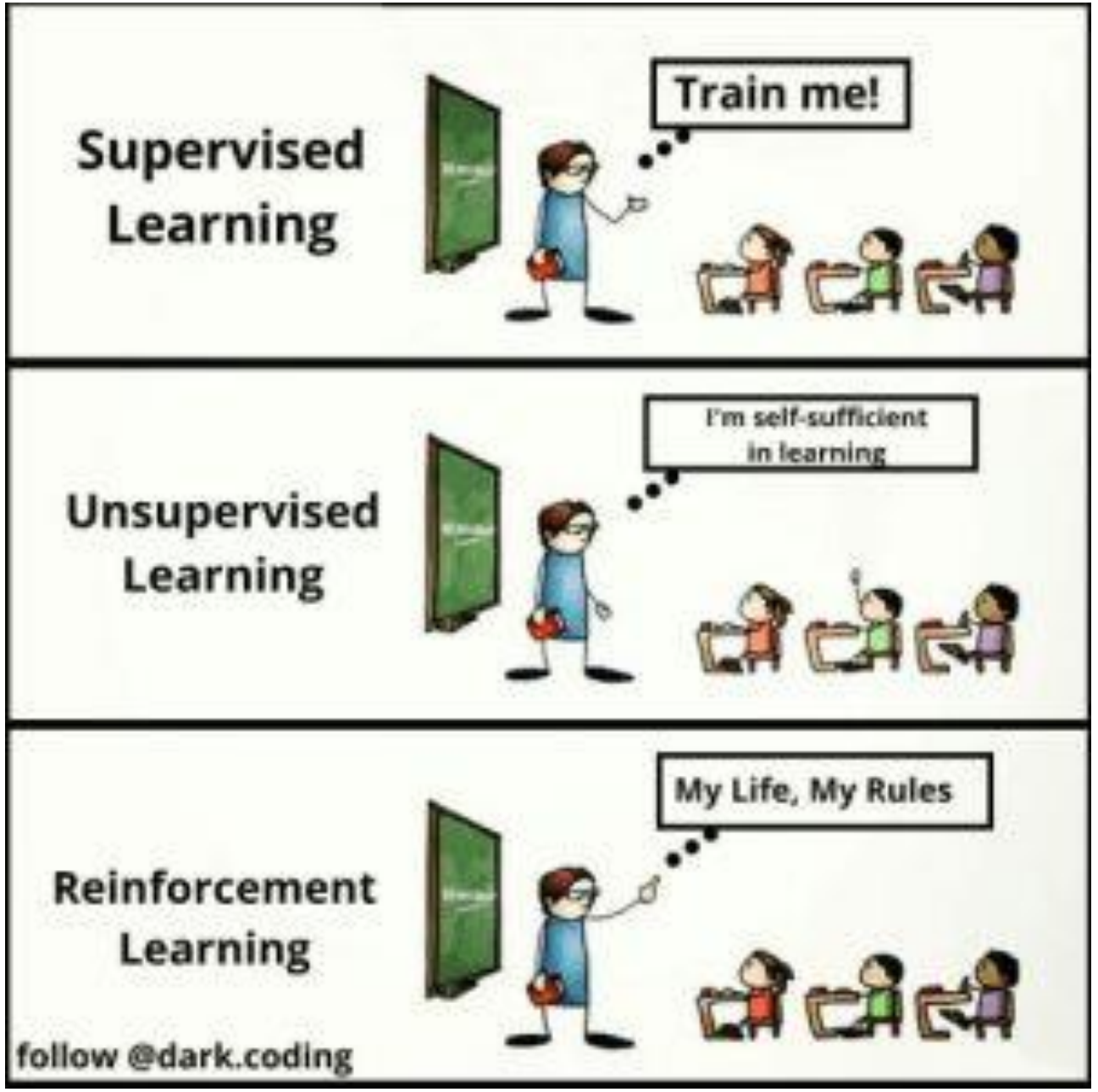
Semi-supervised learning aims to label unlabeled data points using knowledge learned from a small number of labeled data points.

SELF_SUPERVISED LEARNING

REINFORCEMENT LEARNING

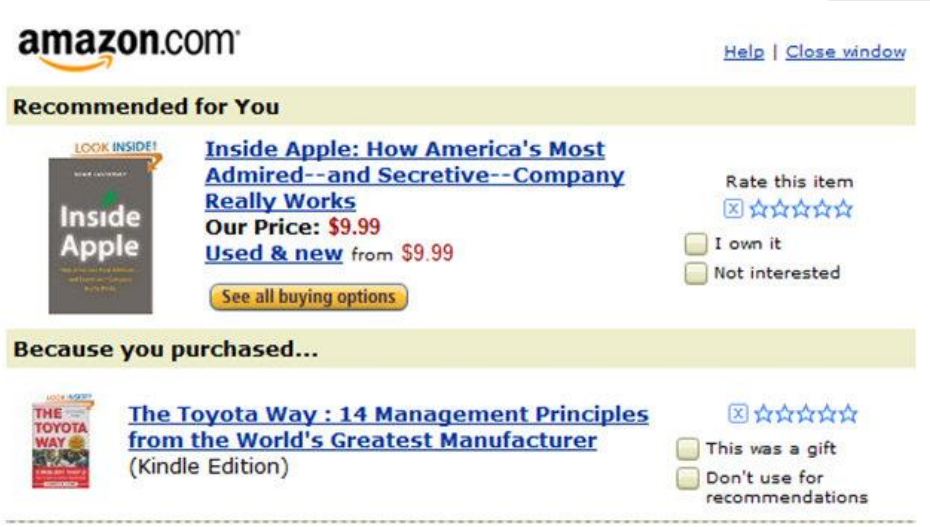
- since the training dataset is missing, it is **bound to learn** from its **experience**

- AlphaZero [Master the games of chess, shogi and go]
- Link: https://www.chess-and-strategy.com/2017/12/intelligence-artificielle-alphazero.html


Supervised Learning
Train me!
Unsupervised Learning
I'm self-sufficient in learning
Reinforcement Learning
My Life, My Rules
follow @dark.coding

AMAZON AUGMENTE SON CA GRÂCE À LA RECOMMANDATION

- Emails personnalisés et recommandations sur site:
  - Contenu "tendance"
  - Articles achetés ensemble
  - Recommandations grâce à l'historique d'achat
  - Recommandations grâce à l'historique des produits vus
  - Nouvelles versions d'un produit déjà possédé



➡ 60% de conversion★, 35% du chiffre d'affaire★★

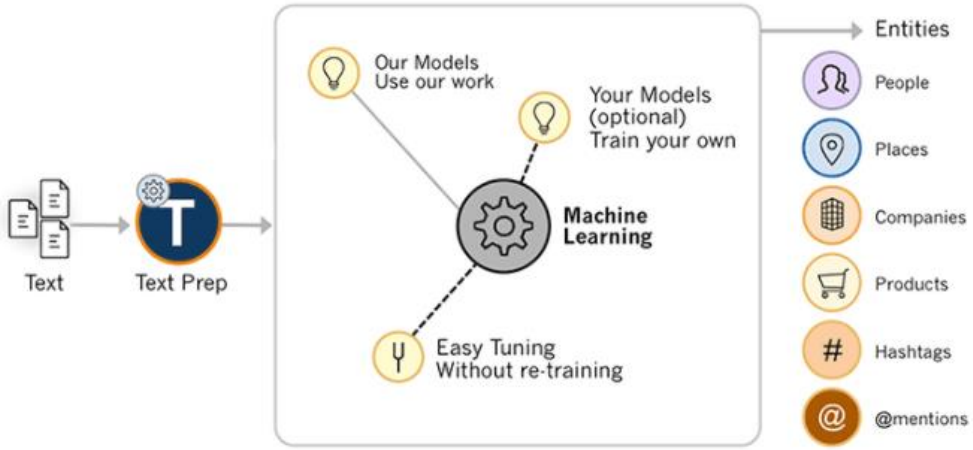
★ Taux de conversion des emails personnalisés tel que communiqué aux actionnaires
★★ Estimation McKinsey en 2013

- Google COLAB : 2 WORKSHOP yelo_V4

PNL : PROGRAMMATION NEURO-LINGUISTIQUE


NLP
Natural lan
processing

NLP


Text
Text Prep
Our Models
Use our work
Your Models
(optional)
Train your own
Machine
Learning
Easy Tuning
Without re-training
Entities
People
Places
Companies
Products
Hashtags
@mentions

NER

```
import spacy
import random


TRAIN_DATA = [('what is the price of polo?', {'entities': [(21, 25, 'PrdName')]}),
              ('what is the price of ball?', {'entities': [(21, 25, 'PrdName')]}),
              ('what is the price of jegging?', {'entities': [(21, 28, 'PrdName')]}),
              ('what is the price of t-shirt?', {'entities': [(21, 28, 'PrdName')]}),
              ('what is the price of jeans?', {'entities': [(21, 26, 'PrdName')]}),
              ('what is the price of bat?', {'entities': [(21, 24, 'PrdName')]}),
              ('what is the price of shirt?', {'entities': [(21, 26, 'PrdName')]}),
              ('what is the price of bag?', {'entities': [(21, 24, 'PrdName')]}),
              ('what is the price of cup?', {'entities': [(21, 24, 'PrdName')]}),
              ('what is the price of jug?', {'entities': [(21, 24, 'PrdName')]}),
              ('what is the price of plate?', {'entities': [(21, 26, 'PrdName')]}),
              ('what is the price of glass?', {'entities': [(21, 26, 'PrdName')]}),
              ('what is the price of moniter?', {'entities': [(21, 28, 'PrdName')]}),
              ('what is the price of desktop?', {'entities': [(21, 28, 'PrdName')]}),
              ('what is the price of bottle?', {'entities': [(21, 27, 'PrdName')]}),
              ('what is the price of mouse?', {'entities': [(21, 26, 'PrdName')]}),
              ('what is the price of keyboad?', {'entities': [(21, 28, 'PrdName')]}),
              ('what is the price of chair?', {'entities': [(21, 26, 'PrdName')]}),
              ('what is the price of table?', {'entities': [(21, 26, 'PrdName')]}),
              ('what is the price of watch?', {'entities': [(21, 26, 'PrdName')]})]
```

- Link= https://manivannan-ai.medium.com/how-to-train-ner-with-custom-training-data-using-spacy-188e0e508c6

SENTIMENT ANALYSIS

- Link:https://github.com/chaymafourati/TUNIZI-Sentiment-Analysis-Tunisian-Arabizi-Dataset/blob/master/TUNIZI-Dataset.txt

```
166  1;taaml aleha kif wlh
167  1;waaaw
168  1;chkoun mazel yasma3 feha
169  1;bravo
170  1;bayrem ya sahbi ya ghali vive la tunisie
```

-1;soooo fake n stupid

-1;techbah lweld houmty

HOW TO TRAIN TEXT CLASSIFICATION MODEL IN SPACY?

```
[('Absolutely wonderful - silky and sexy and comfortable', 1),
('Love this dress!  it\'s sooo pretty.  i happened to find it in a store, and i\'m gl
 1),
('I had such high hopes for this dress and really wanted it to work for me. i initial
 0),
("I love, love, love this jumpsuit. it's fun, flirty, and fabulous! every time i wear
 1),
('This shirt is very flattering to all due to the adjustable front tie. it is the per
 1),
('I love tracy reese dresses, but this one is not for the very petite. i am just unde
 0),
('I aded this in my basket at hte last mintue to see what it would look like in persc
 1),
("I ordered this in carbon for store pick up, and had a ton of stuff (as always) to t
 1),
('I love this dress. i usually get an xs but it runs a little snug in bust so i order
 1),
('I\'m 5"5\' and 125 lbs. i ordered the s petite to make sure the length wasn\'t too
 1)]
```

LINK:HTTPS://OPENAI.COM/BLOG/DALL-E/

DALL·E[1] is a 12-billion parameter version of GPT-3 trained to generate images from text descriptions, using a dataset of text–image pairs

- Link:
- https://github.com/taghouti-ghofrane/Creative_Lab_EVENTS/blob/main/Creative_Lab_Event_Workshop_By_Taghouti_Ghofrane.ipynb

- https://www.machinelearningplus.com/nlp/custom-text-classification-spacy/

COMPETITION

- Quiz (10 Points)

Project : (10 Points)

Website Link : https://www.automobile.tn/fr

1- Scrap The comments of The Users of the website (Automobile.tn) (5 Points)

2_ Create a Model for Sentiment Analysis (TUNIZI-Sentiment-Analysis) (5 points)

Or

3_ Try To be Smart and find Your own solution To make me See something Realted to Sentiment analysis (Ps: Language is Your secret ^^)

Good Luck