FSD322 - GI / IF

Jonathan Sprauel

Planning of the module

Tuesday 28th	Wednesday 29th
8h-12h30 : Introduction Vocabulary, overview of concepts and problems	*8h-12h30 :* TP : Low Code app using Dataiku
13h30-18h : Data Analysis : Jupyter, pandas, sklearn, visualization	*13h45-17h :* Advanced applications Data Lifecycle

\+ Evaluation quizz at end of module

Part 1 : Introduction to AI

Tuesday 28th	Wednesday 29th
8h-12h30 : Introduction Vocabulary, overview of concepts and problems	*8h-12h30 :* TP : Low Code app using Dataiku
13h30-18h : Data Analysis : Jupyter, pandas, sklearn, visualization	*13h45-17h :* Advanced applications Data Lifecycle

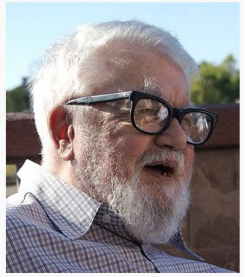
- Basic definitions of AI and its applications
- Use cases in industrial engineering (process optimization, predictive maintenance) and finance (predictive analysis, automated trading)
- Case studies: improving productivity through AI, using AI in portfolio management

UNE DÉFINITION DE L'IA ?

WIKIPÉDIA
L'encyclopédie libre

L'**intelligence artificielle** (**IA**) est un « ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine »

« La construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains »



John McCarthy
Pionnier de l'IA avec M.L Minsky

Programmes qui résolvent des tâches **compliquées** : celles qui ne sont accomplies aujourd'hui que par **l'être humain**

The different types of learning

Supervised Learning

- Learning with a **labeled** training set.

Learn with exercises
Ex. Driving license

Unsupervised Learning

- Discovering patterns in **unlabeled** data.

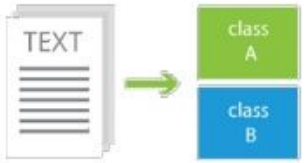
Learn with similitude
Ex. Newton and the apple

Reinforcement Learning

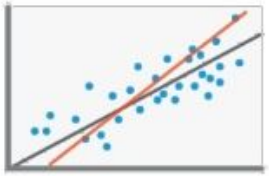
- Learning based on **feedback** or **reward**.

Learn with trial and error
Ex. Ride a bike

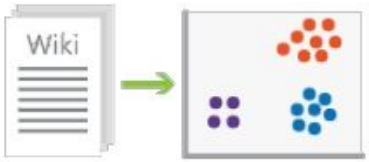
ML to solve different types of problems



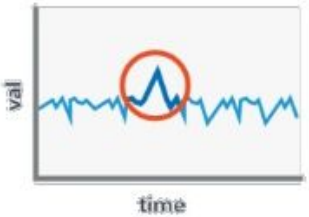
Classification
(supervised – predictive)



Regression
(supervised – predictive)



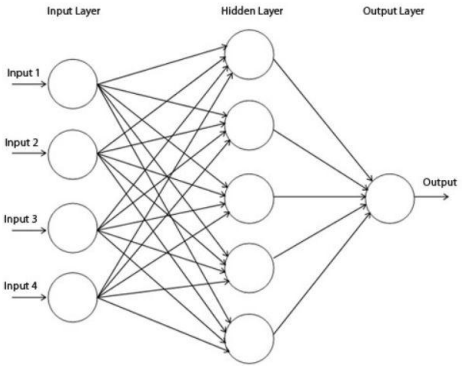
Clustering
(unsupervised – descriptive)



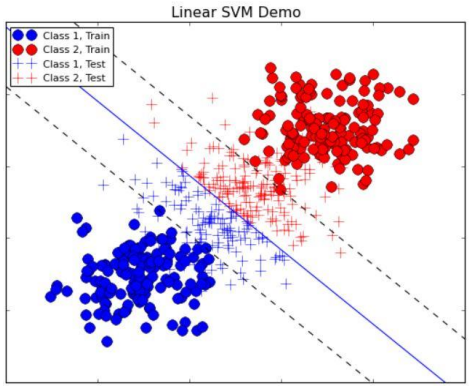
Anomaly Detection
(unsupervised – descriptive)

Classical Machine Learning

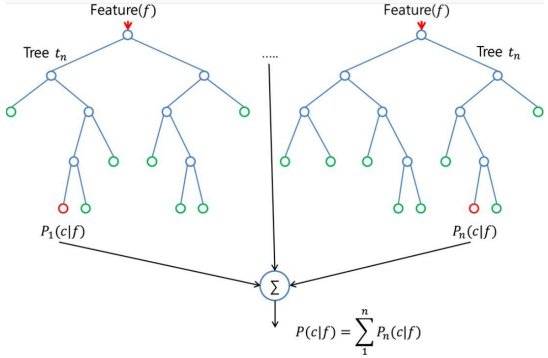
Multi-Layer Perceptron (1986)



SVM (1995)



Random forest (2001)



A brief history of Deep Learning

- **1950** • Test de Turing
- **1981** • Fukushima Neocognitron : lecture d'écriture manuscrite en Japonais
- **1988:** • Convolutional Network (**CNN**) de LeCun lecture d'adresse postale. 60k paramètres
- **2012** • Traffic Signs Challenge : Performances meilleures que les humains. AlexNet : 60 M paramètres
- **2016** • Alphago bat le champion du monde de go.
- **2023** • GPT4 : 175 Milliards de paramètres

Google PaLM 2 : 540 Milliards de paramètres

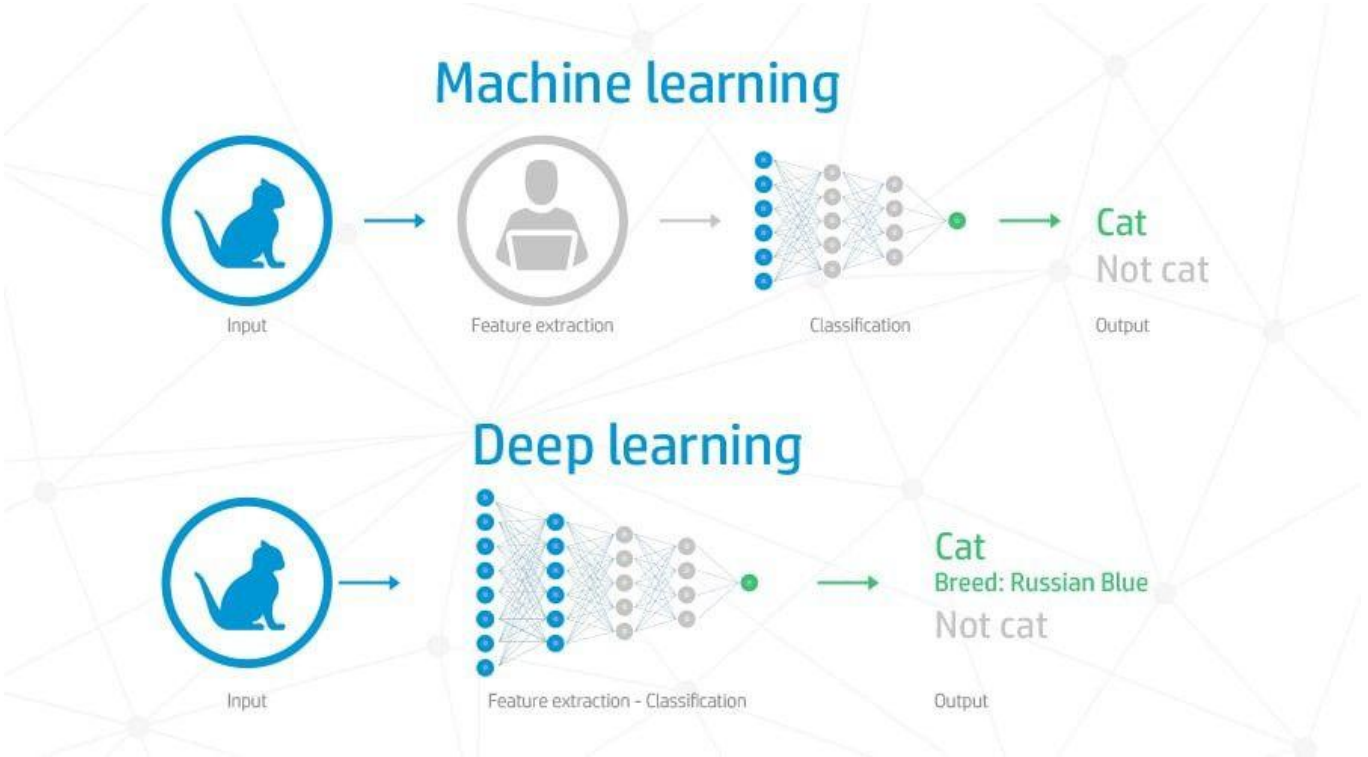
MIT Technology Review

Facebook Launches Advanced AI Effort to Find Meaning in Your Posts

A technique called deep learning could help Facebook understand its users and their data better.

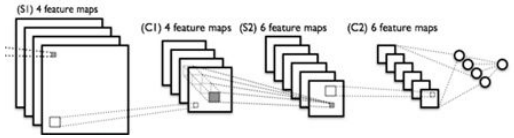
© reuters/ Kim Hong Ji

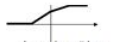
Machine Learning != Deep Learning != Artificial Intelligence



A brief history of what happened

New models:
convolutions,
subsampling, etc



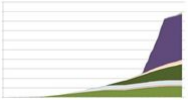


MLP 1986

Tricks on model architectures,
Learning methods

High performance computing,
GPUs for training models

Big Data : a lot of data available for training models



Open source community very active

Buzz cleverly orchestrated
(Google, Facebook, etc.)

Deep Learning

Exercice 1 : Regression

Objectives :

- Understand the difference between Regression and Classification

- Understand the definition of a Label

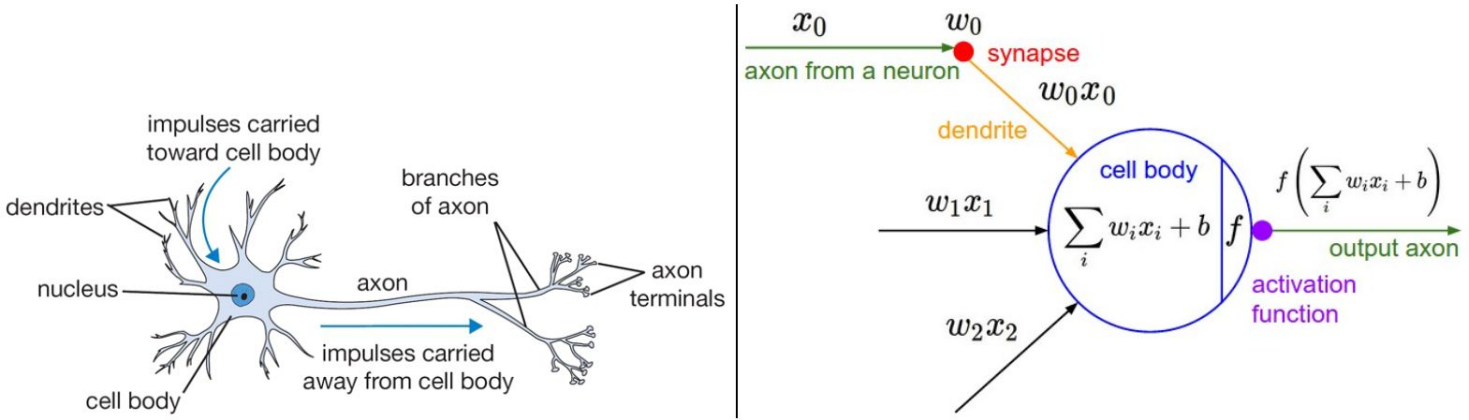
Exercice 2 : Features

Objectives :

- Understand the notion of Feature
- Understand the importance of Feature selection
- Understand how Deep learning changes the computation of Features

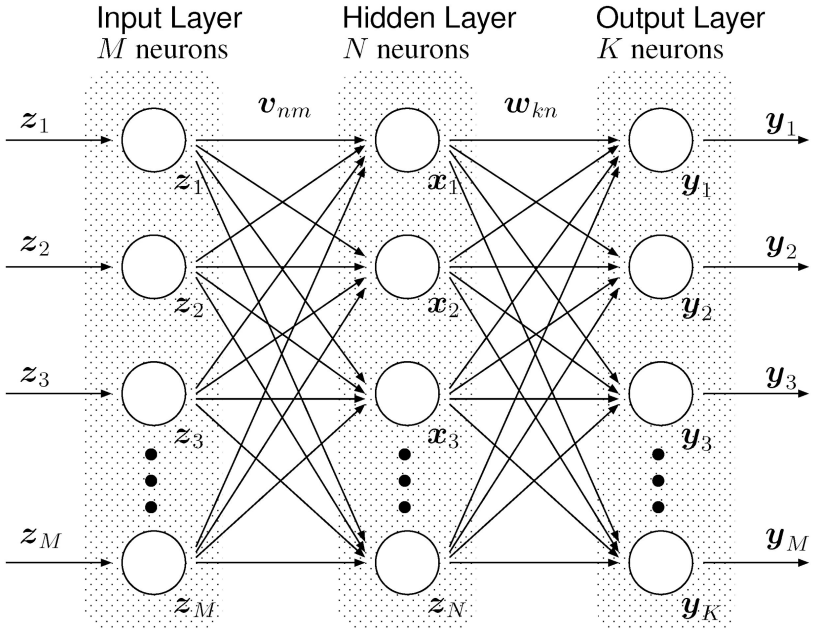
Neurons

- Neurons are trained to filter and detect features such as edges, shapes, textures, by receiving weighted inputs from the previous neurons, transforming it with an activation function and passing it to the outgoing connections.



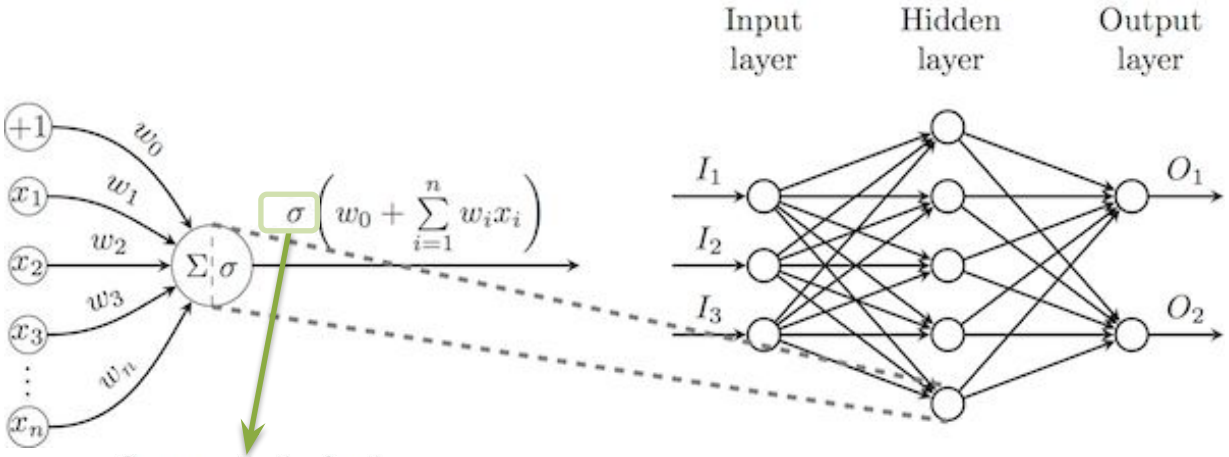
Multi-layer Perceptron (MLP)

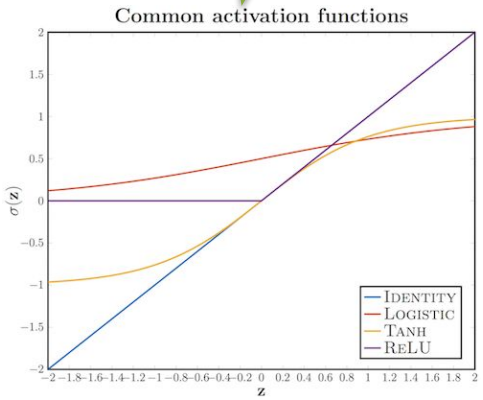
- MLP interest is in the association of neurons in multi layers : it results in a composition of non linear functions that can represent complex problematics.



- Parameters estimation:
- Quadratic error is known (estimated – known)² => we can estimate the gradient for the last layer
- We don't know the quadratic error associated to each hidden layer.

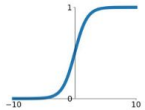
Activation Functions





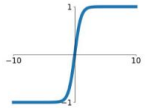
Sigmoid

$\sigma(x) = \frac{1}{1+e^{-x}}$



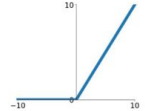
tanh

$\tanh(x)$



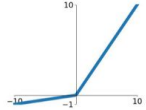
ReLU

$\max(0, x)$



Leaky ReLU

$\max(0.1x, x)$

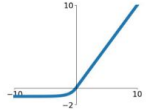


Maxout

$\max(w_1^T x + b_1, w_2^T x + b_2)$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



Exercice 3 : Neurones

Objectives :

- Understand the influence of hyper-parameters
- Reinforce the notion of Feature and the distinction between ML and DL

playground.tensorflow.org/

Quizz time : Fill in the definitions

Level 1	Machine Learning	Deep Learning	Artificial Intelligence	Big Data
Level 2	Supervised vs Unsupervised learning	Classification vs Regression	Correlation	Feature vs target
Level 3	Overfitting	Hyper parameter	Training vs Testing Dataset	Feature engineering

Quizz time : Some answers

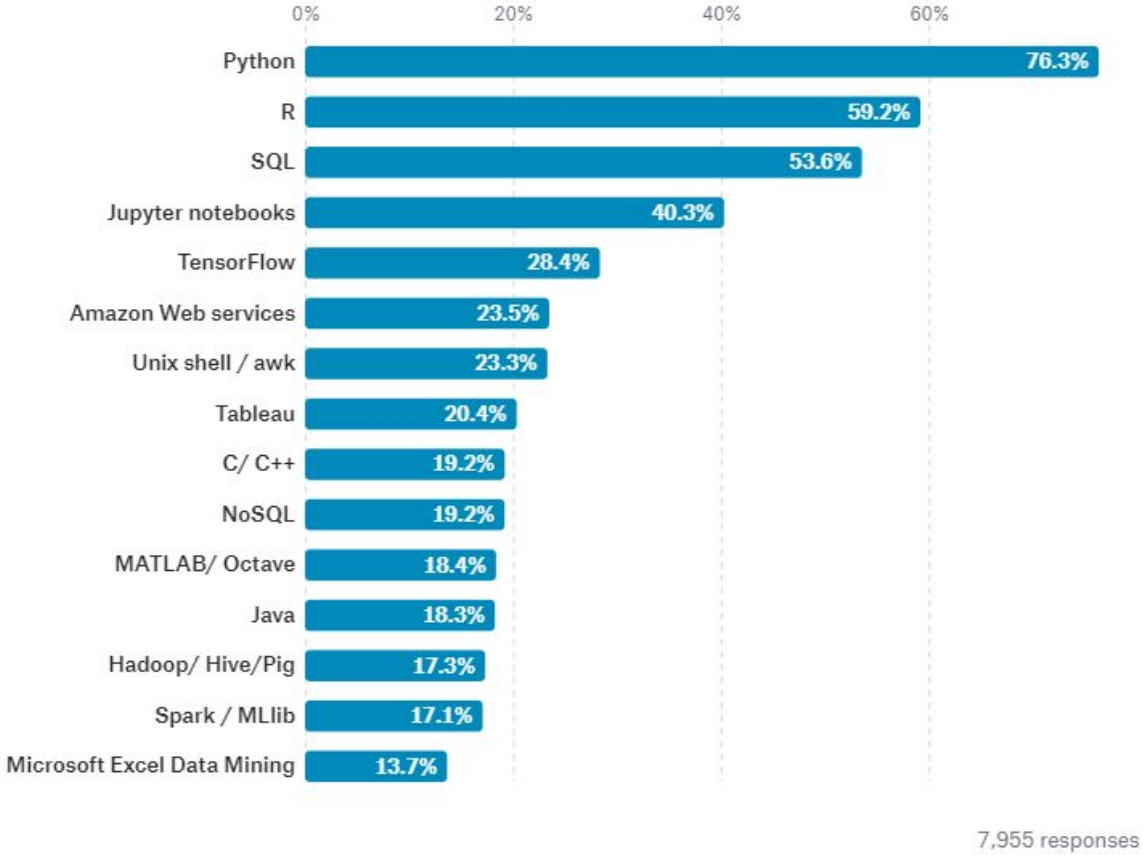
Machine learning is a field of computer science that gives computer systems the ability to “learn” (i.e. progressively improve performance on a specific task) with data, without being explicitly programmed. (Wikipedia)

Artificial intelligence (AI) is the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. (Brittanica)

Big Data refers to working with datasets that have large Volume,Variety, Velocity (, Veracity, and Value).

Deep Learning is Machine Learning with Deep Neural Networks.

Which tools are used



What should I look for in a data scientist's CV?

Must have :

- Technology names (most of them) : sklearn, python / R, keras / tensorflow, jupyter, numpy, pandas, spark
- Experiences with datasets outside of a MOOC
- Likes understanding people's problems

Nice to have :

- PhD (in computer science, applied math or physics)
- kaggle competition/score
- publications (Arxiv, JMLR, MLJ, IEEE PAMI, NIPS, ICML, ICLR…)
- cloud experience (AWS,GCP, Azure) or deployment experience (docker, terraform, kubernetes,...)

Sklearn : lets have a look

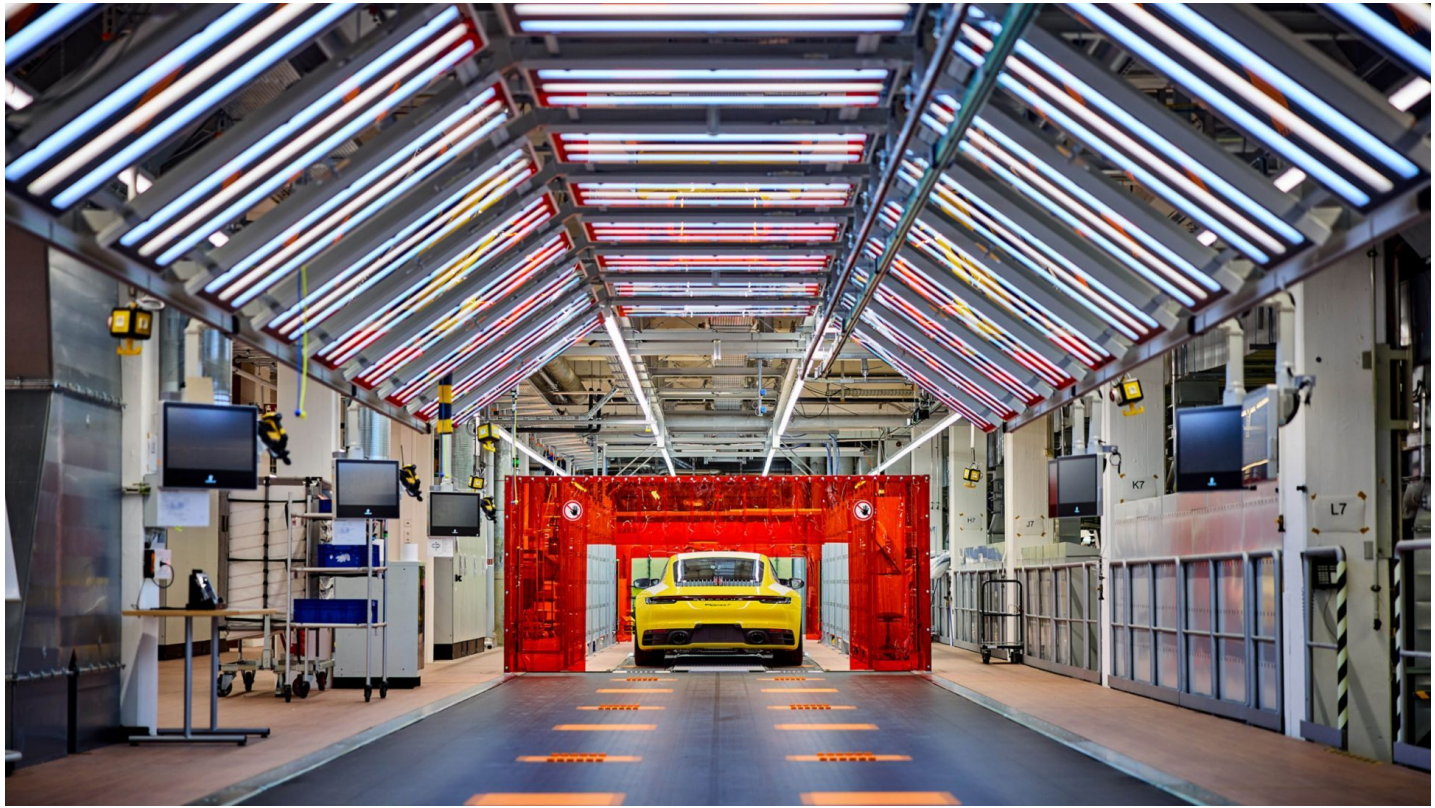
http://scikit-learn.org

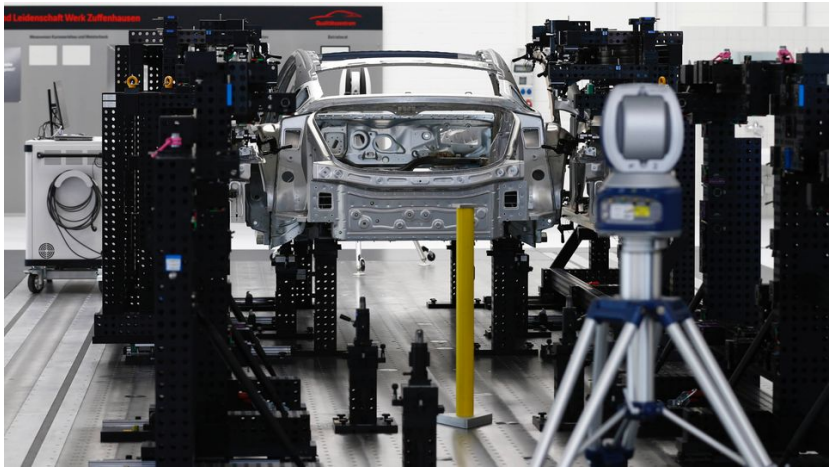
AI in Process Optimization for Industrial Engineering

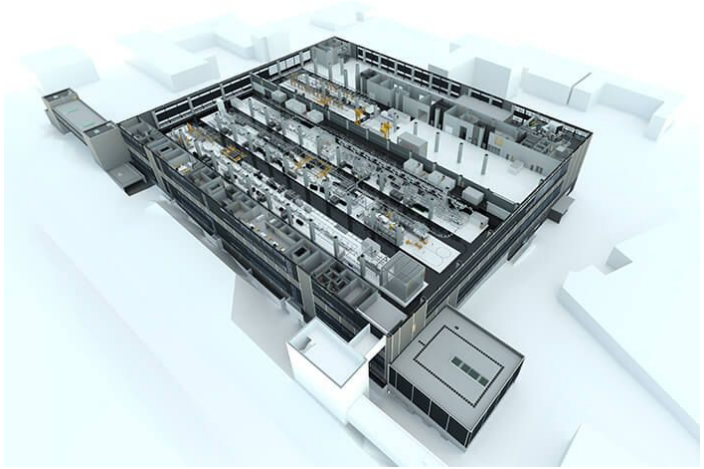
Role of AI in Process Optimization:

- **Enhanced Decision-Making**: AI algorithms analyze vast datasets to make data-driven decisions, leading to optimized processes.
- **Real-time Monitoring**: AI systems provide real-time monitoring of various parameters, allowing for quick adjustments and preventive measures.
- **Optimizing Production Lines**: AI identifies bottlenecks, predicts equipment failures, and streamlines production workflows for increased efficiency.
- **Predictive Maintenance**: Anticipating equipment maintenance needs through AI reduces downtime and extends the lifespan of machinery.
- **Supply Chain Optimization**: AI algorithms improve logistics, demand forecasting, and inventory management, leading to cost savings.

Reference Use Case : Porsche Factory 4.0







AGV - Cobotic

Automated Checkpoint

Digital Twin

Product line forecasting and optimization

AI for Predictive Analysis and Automated Trading

- **Predictive analysis** utilizes AI algorithms to analyze historical data, identify patterns, and make predictions about future market trends.
  - Market Trend Prediction: AI algorithms analyze market indicators, news, and historical data to predict future market trends with higher accuracy.
  - Risk Management: Predictive models assess potential risks by analyzing market volatility, economic indicators, and external factors, aiding in informed decision-making.

- **Automated trading** involves using AI algorithms to execute buy or sell orders on financial markets without human intervention.
  - Hedge funds outperformance compared to traditional strategies.
  - Reduced Latency: Investment firms adopting AI-driven automated trading experience significantly reduced latency, resulting in improved execution times.

But also : information summary, e-commerce automations, fraud and bad actors detections, biometrics, …

Part 2 : Data Analysis (aka replace Excel)

Tuesday 28th	Wednesday 29th
8h-12h30 : Introduction Vocabulary, overview of concepts and problems	*8h-12h30 :* TP : Low Code app using Dataiku
13h30-18h : Data Analysis : Jupyter, pandas, sklearn, visualization	*13h45-17h :* Advanced applications Data Lifecycle

- Collection and preprocessing of data
- Basic concepts of supervised and unsupervised learning
- Exploration of popular algorithms (clustering, regression, classification)
- Techniques of feature engineering for AI models
- Practical case: cleaning and preparation of financial data

https://github.com/a3lentyr/machine-learning-short

Part 3 : Dataiku

Tuesday 28th	Wednesday 29th
8h-12h30 : Introduction Vocabulary, overview of concepts and problems	*8h-12h30 :* TP : Low Code app using Dataiku
13h30-18h : Data Analysis : Jupyter, pandas, sklearn, visualization	*13h45-17h :* Advanced applications Data Lifecycle

- Using Dataiku

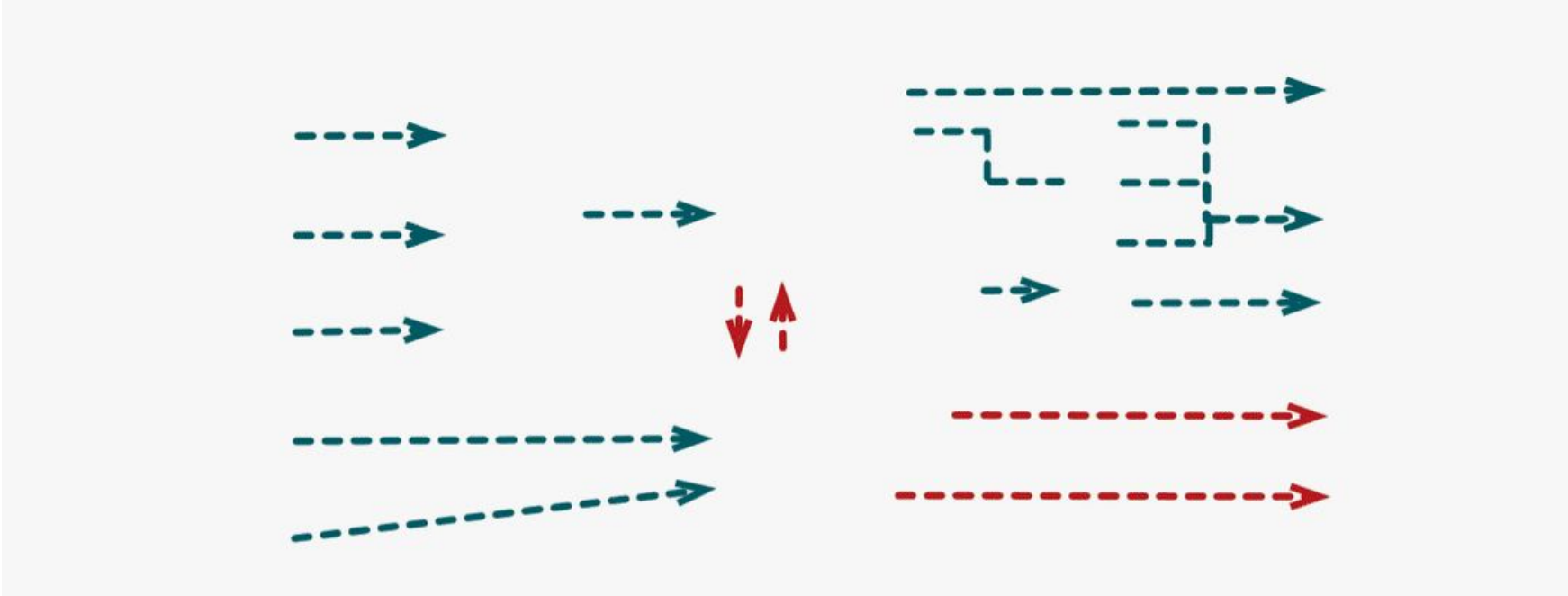
Part 4 : Advanced Applications

Tuesday 28th	Wednesday 29th
8h-12h30 : Introduction Vocabulary, overview of concepts and problems	*8h-12h30 :* TP : Low Code app using Dataiku
13h30-18h : Data Analysis : Jupyter, pandas, sklearn, visualization	*13h45-17h :* Advanced applications Data Lifecycle

- Data Lifecycle
- Use of sensors and IoT for data acquisition

- Introduction to neural networks
- Case study: prediction models (financial forecasting)

Data Lifecycle



Data Lifecycle

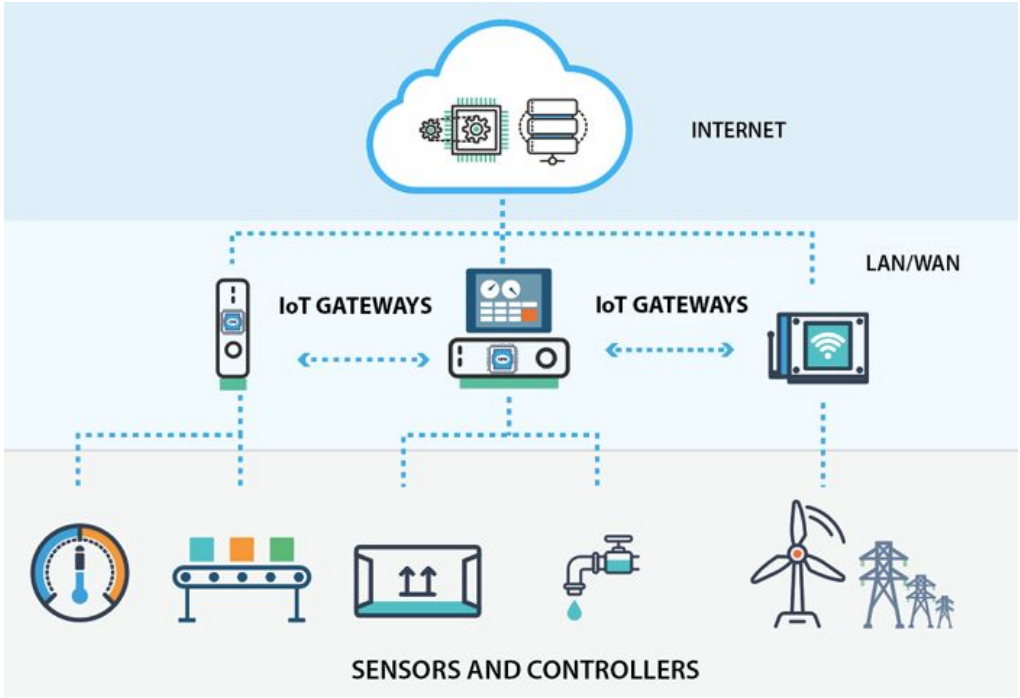
ETL (Extract, Transform, Load) is a process that involves extracting data from source systems, transforming it into a usable format, and loading it into a target system

A **Data Lake** is a storage repository that holds vast amounts of raw data in its native format until it is needed.

A **Data Warehouse** is a centralized repository that consolidates structured data from various sources for reporting and analysis.
It is optimized for query and analysis and often involves data cleaning, integration, and transformation.

Data Marts are subsets of a data warehouse that focus on specific business lines, departments, or user groups. They are designed to meet the specific needs of a particular group

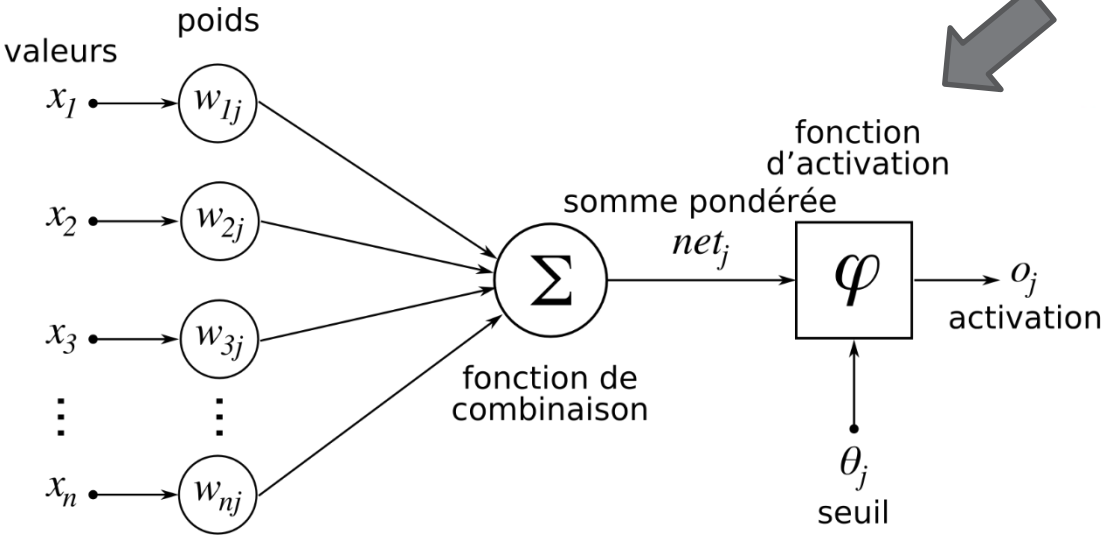
Use of sensors and IoT for data acquisition in Factory 4.0

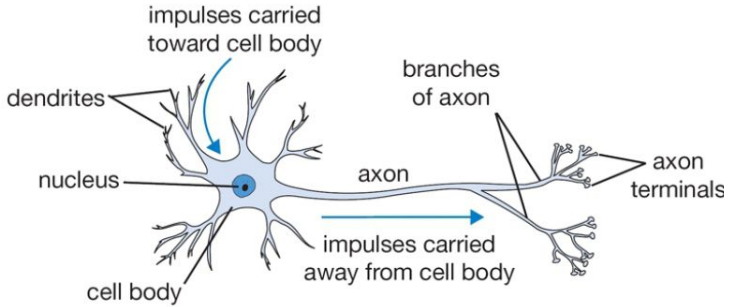


- A network of sensors and devices throughout the factory floor to collect real-time data on machines, equipment, and processes.
- Process and analyze the vast amount of data generated by IoT devices
- Predictive maintenance strategies based on IoT data to anticipate equipment failures, minimize downtime, and optimize maintenance schedules.
- Remote Monitoring and Control of machines and processes, allowing for real-time control and intervention, even from a centralized control room.

ET LES RÉSEAUX DE NEURONES ?

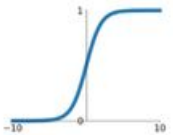
Une fonction mathématique **non linéaire** inspirée des neurones biologiques
Elle contient des **poids** qui sont réglés par un **entrainement**.





Sigmoid

$$\varphi(x) = \frac{1}{1+e^{-x}}$$

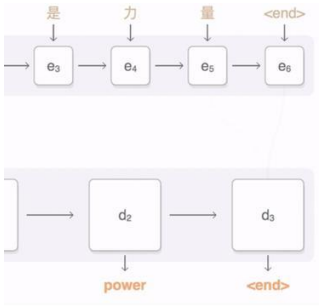


LES APPLICATIONS *RÉSOLUES*

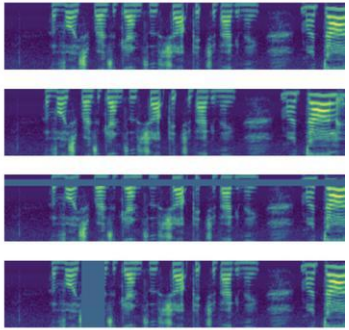
Image Classification :
91% on Image Net

Object Detection

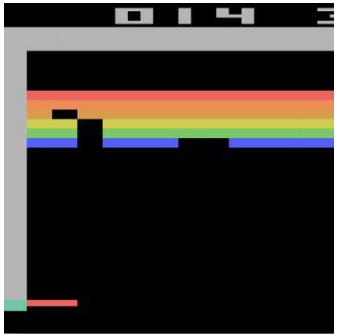
Analyse de commentaires
(amazon, twitter, ...)
96% on IMDB



Machine Translation
BLEU score 40
(34 human pro)



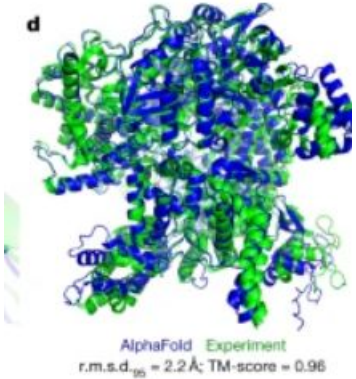
Speech Recognition
97% on Noisy



Atari, Chess, Go

LES APPLICATIONS QUI SONT ENCORE DANS LE DOMAINE DE LA RECHERCHE

Image & Video Generation
Diffusion Models



Protein Prediction
>90% AlphaFold 2

Conversation agents (LLM)
67,6% Accuracy on US Medical exam

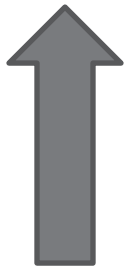
Jeux multi agents *:*
Starcraft, Diplomacy…

DES FRAMEWORK DANS CHAQUE LANGAGE

Caffe · Chainer · DL4J Deeplearning4j · KERAS · Microsoft CNTK · MatConvNet

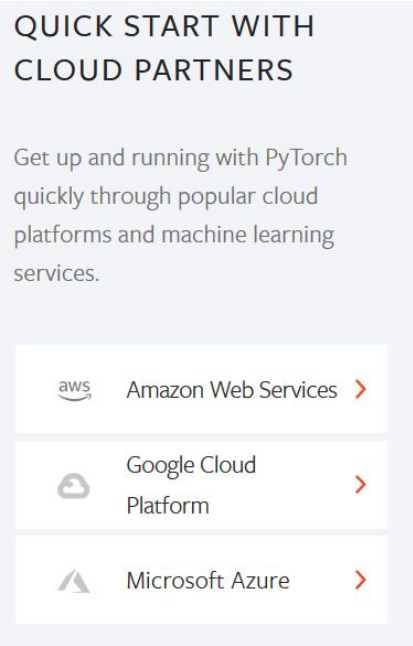
MINERVA · mxnet · Purine · TensorFlow · theano · torch

```
class MNISTModel(LightningModule):
    def __init__(self):
        super().__init__()
        self.l1 = torch.nn.Linear(28 * 28, 10)

    def forward(self, x):
        return torch.relu(self.l1(x.view(x.size(0), -1)))

    def training_step(self, batch, batch_nb):
        x, y = batch
        loss = F.cross_entropy(self(x), y)
        return loss

    def configure_optimizers(self):
        return torch.optim.Adam(self.parameters(), lr=0.02)
```



PyTorch en python

ET POUR LE METTRE EN PRODUCTION



https://pytorch.org/

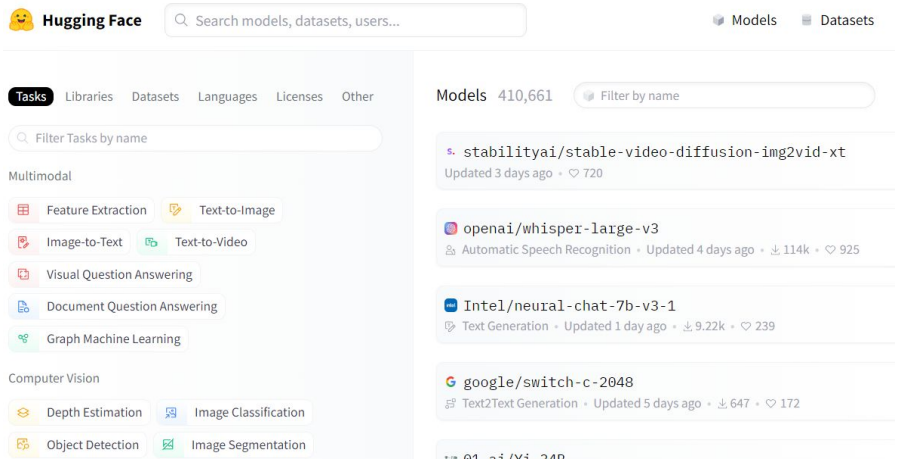
Très facile d'utilisation sur le cloud
Matériel optimisé pour entrainement rapide

Mais…
Un coût qui dépend du nombre de paramètres

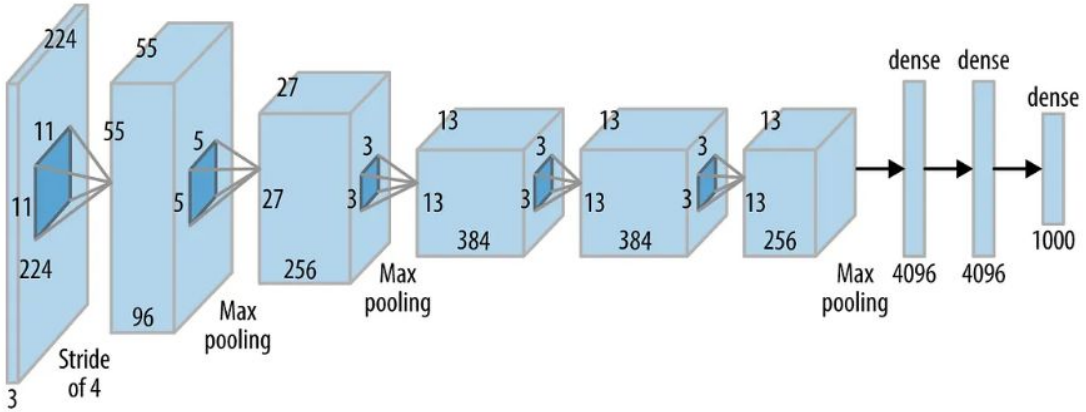
	Cout d'entrainement
DeepMind AlphaGO	35 Million $
GPT3	12 Million $
CoAtNet (top 1 ImageNet)	250 000 $
BERT	7000 $
Yolo V5	100 $
ResNet 50	10 $

PARTIR D'UN MODÈLE DÉJÀ ENTRAINÉ

Hugging Face / Pytorch hub : des modèles de l'état de l'art avec les poids déjà réglés
=> On ajoute des images, on continue l'entrainement, et hop !



Les articles et blogs décrivent des architectures (combien de couches mettre et de quels types), qui sont connus pour bien marcher sur certains problèmes



Alexnet Block Diagram (source:oreilly.com)

Quelques mots clés du domaine

NN: Neural Network
CNN: Convolutional Neural Network
RNN: Recurrent Neural Network
LSTM: Long Short-Term Memory
GAN: Generative Adversarial Network
VAE: Variational Autoencoder

LLM : Large Language Models
ReLU: Rectified Linear Unit
NLP: Natural Language Processing
GPU: Graphics Processing Unit
MLOps: Machine Learning Operations
XAI: Explainable AI