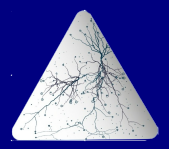
Introduction to Machine Learning and Artificial Intelligence

Hamza BA-MOHAMMED

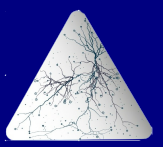
EAIC Vice President charged of Training

November 10, 2022



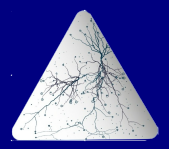
Summary

1. About Machine Learning
2. About Artificial Intelligence
3. AI vs ML: what's the difference?
4. Philosophical discussion about the learning concept and the Theory of Machine Learning
5. Use cases of ML
6. Ethics of AI: more of philosophy
7. Roadmap of the Training Cell
8. A first test: Google AI Teachable Machine
9. Get your hands dirty!
   - My First Machine Learning Code



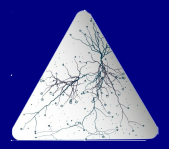
Definition

Machine Learning
the process of making the machine able to **autonomously** learn from the data it got, thus without the intervention of human beings.



Definition

$$\text{data from the past} + \text{parameters} \Longrightarrow \text{algorithms} \Longrightarrow \text{prediction for the future}$$



Types of Machine Learning

Supervised Learning

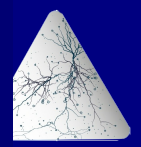
try to model relationships and dependencies between the target prediction output and the input features such that we can predict the output values for new data based on those relationships which it learned from the previous data sets $\Longrightarrow$ **Predictive Model + Labeled data**

Unsupervised Learning

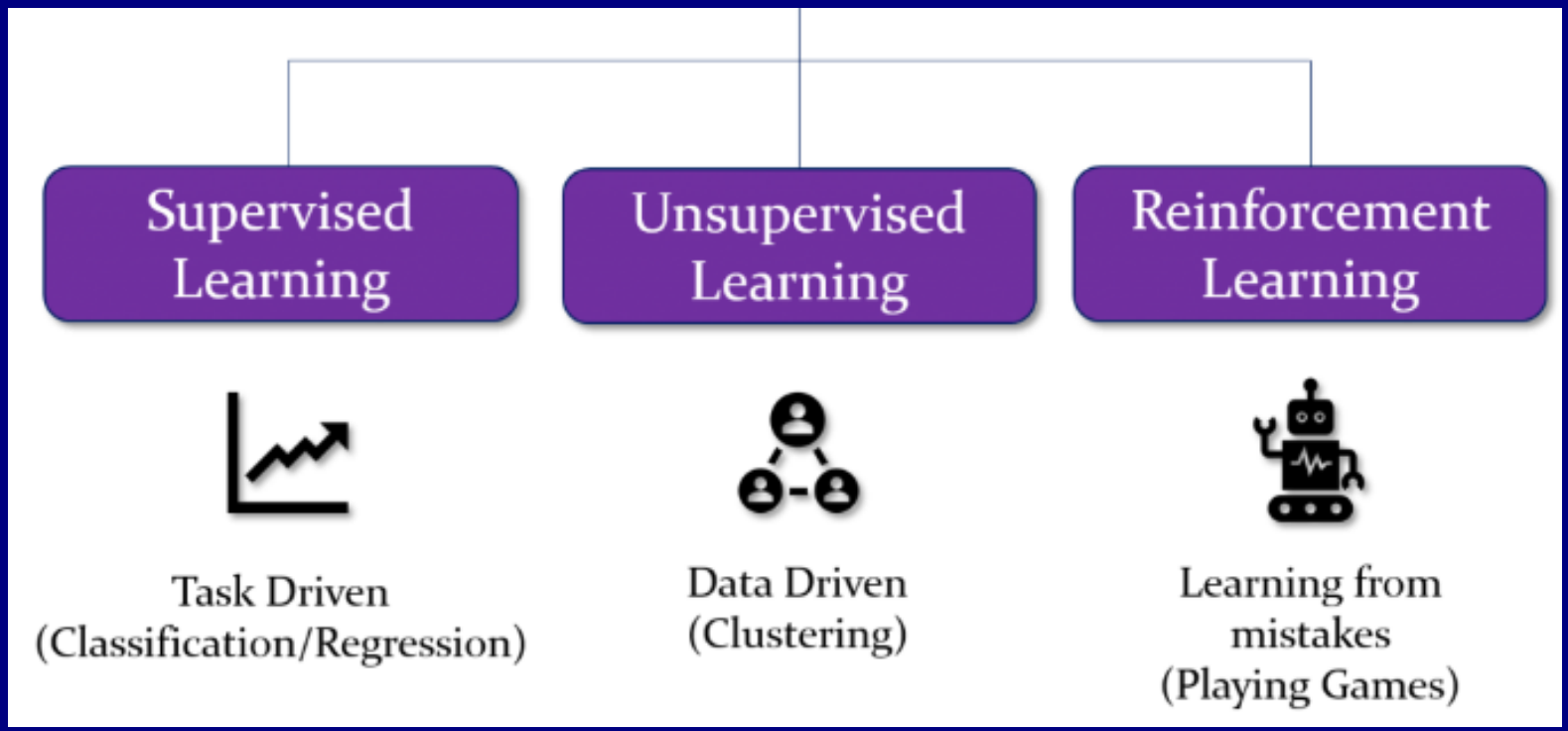
try to use techniques on the input data to mine for rules, detect patterns, summarize and group the data points which help in deriving meaningful insights and describe the data better $\Longrightarrow$ **Descriptive Model + Unlabeled data**

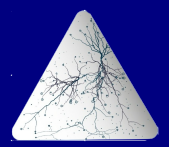
Reinforcement Learning

It allows machines to automatically determine the ideal behavior within a specific context, in order to maximize its performance. Simple reward feedback is required for the agent to learn its behavior (the reinforcement signal) $\Longrightarrow$ **Decisive Model (reward-based) + Labeled data**

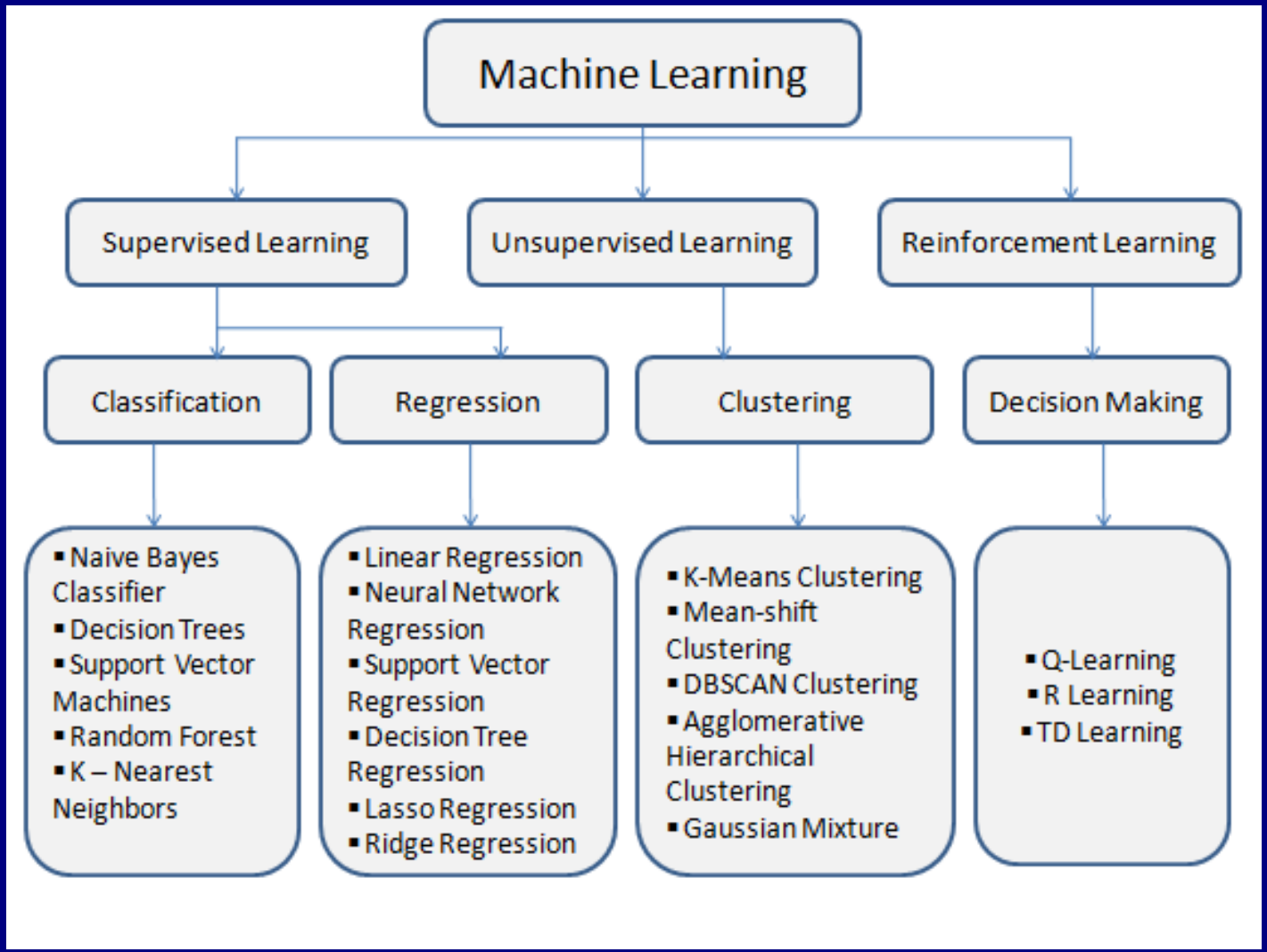


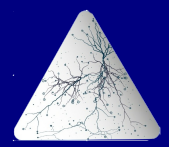
Types of Machine Learning





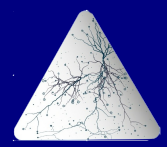
Types of Machine Learning



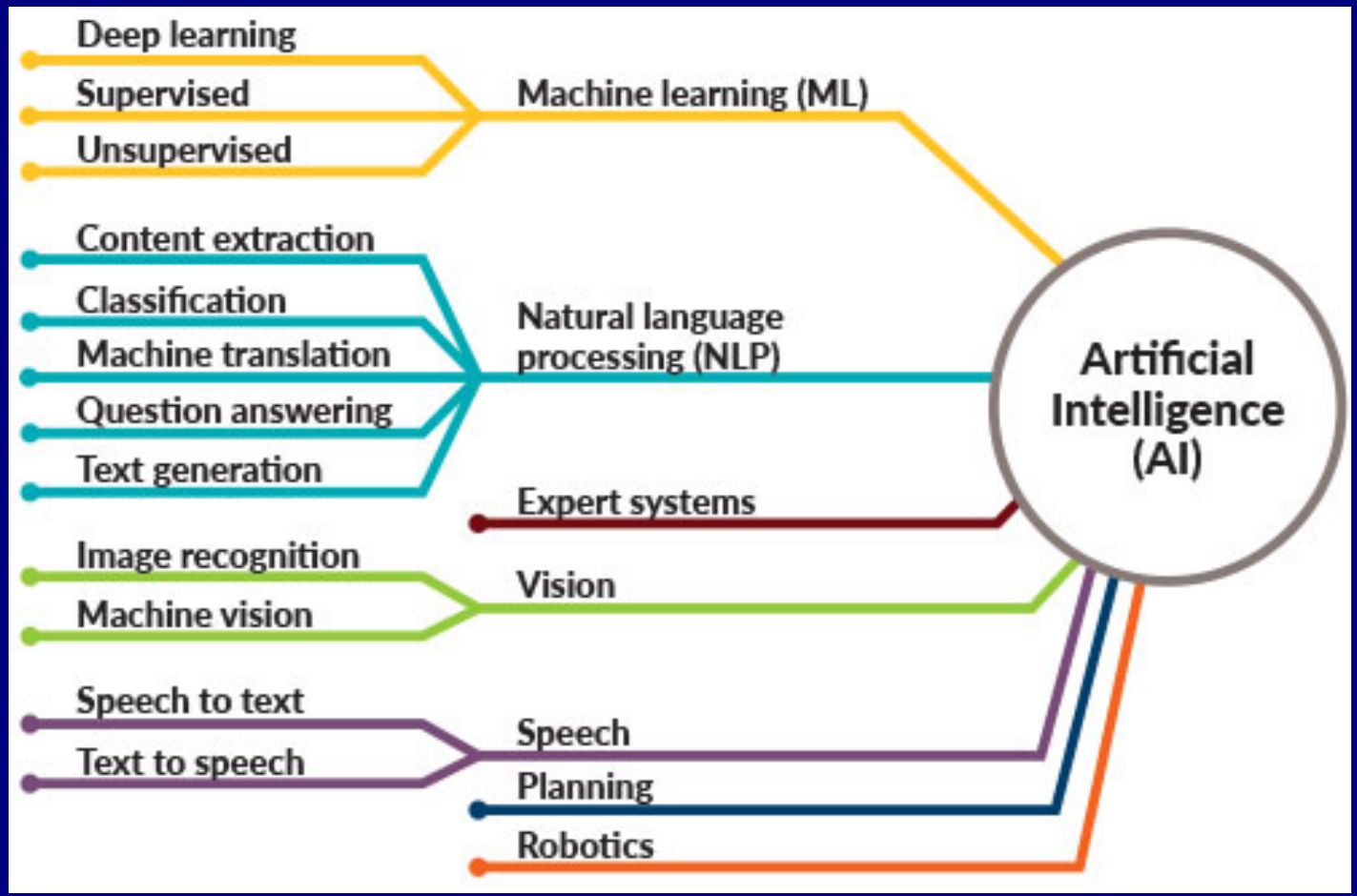


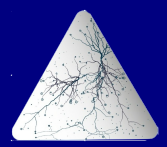
Artificial Intelligence

the process of making the machine able to learn **similarly** to human beings through multiple algorithms.

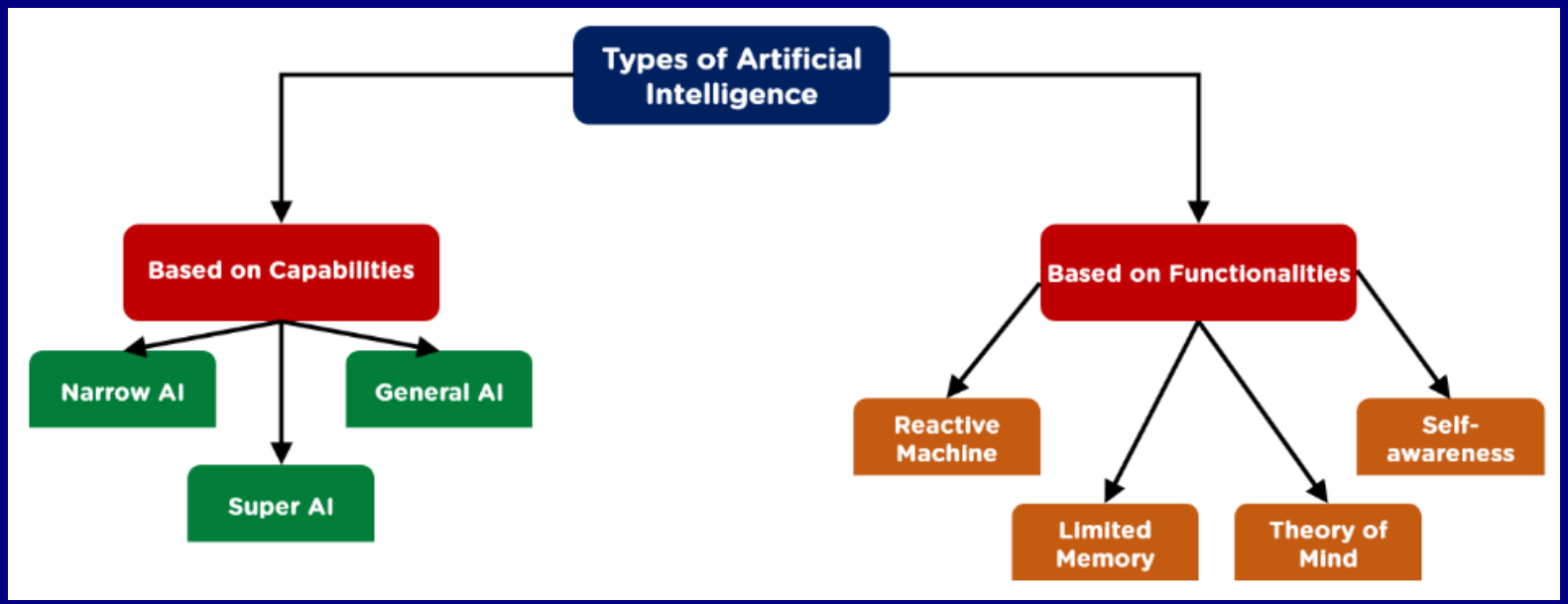


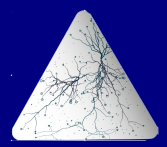
Types of AI


Deep learning
Supervised
Unsupervised
Machine learning (ML)
Content extraction
Classification
Machine translation
Question answering
Text generation
Natural language processing (NLP)
Artificial Intelligence (AI)
Expert systems
Image recognition
Machine vision
Vision
Speech to text
Text to speech
Speech
Planning
Robotics



Types of AI

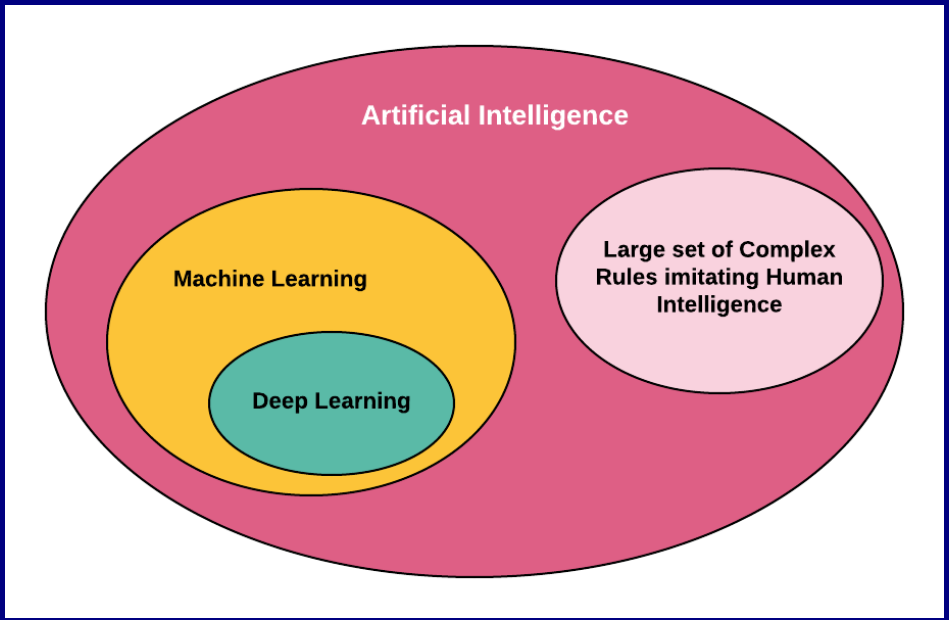

Types of Artificial Intelligence
Based on Capabilities
Narrow AI
General AI
Super AI
Based on Functionalities
Reactive Machine
Limited Memory
Theory of Mind
Self-awareness

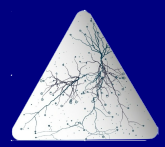


AI is the ocean, ML is a golf

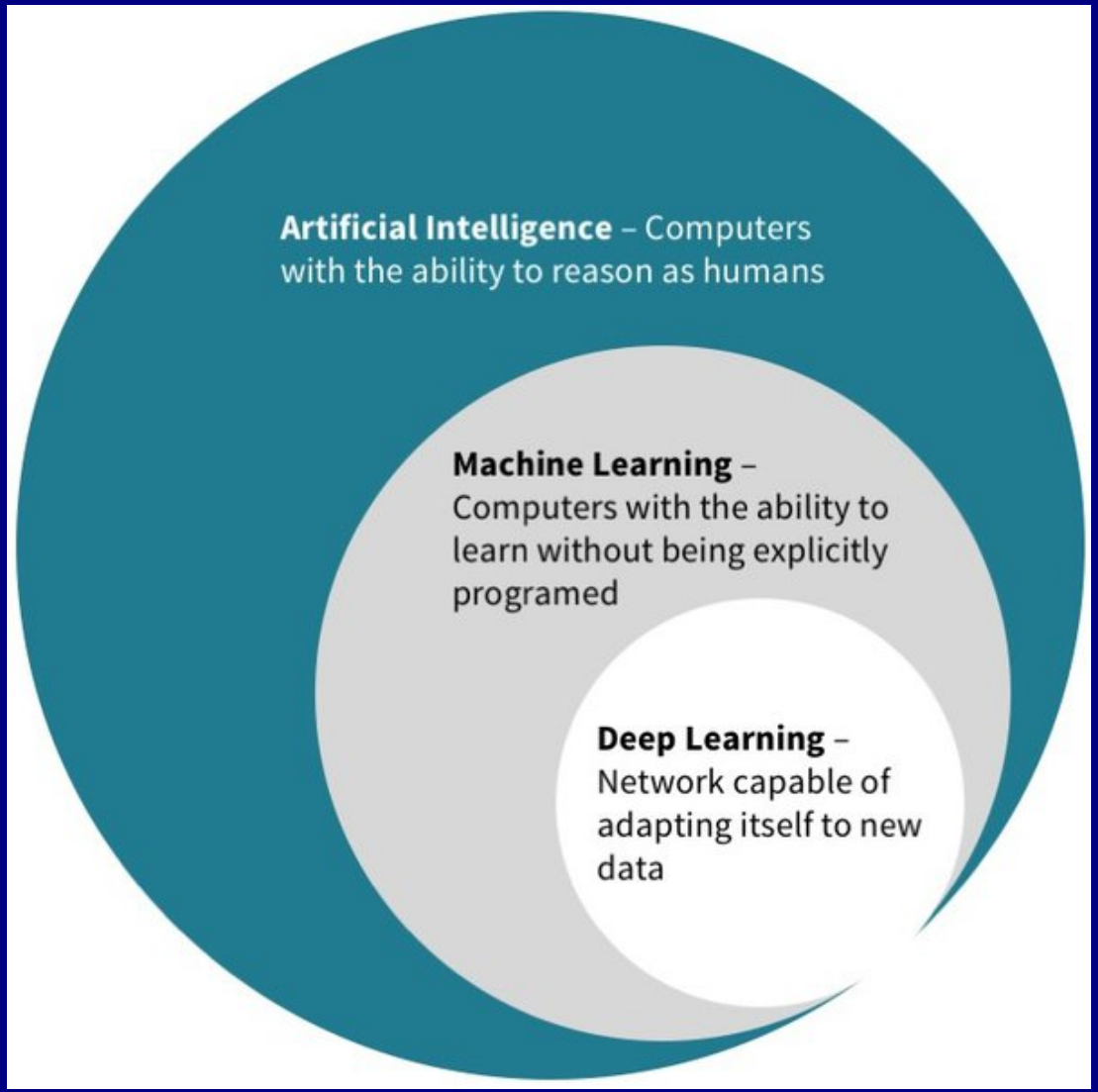
To sum things up..

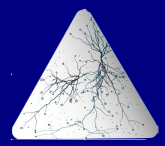
AI solves tasks that require human intelligence while ML is a subset of AI that solves specific tasks by learning from data and making predictions.
This means that all ML is AI, but not all AI is ML.





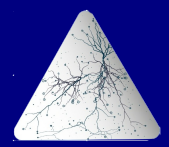
AI is the ocean, ML is a golf





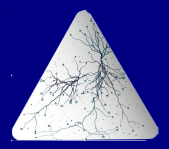
Good and Bad Learning: Philosophy time!

what is "learning"?
how can we identify if a machine learned?
can we evaluate learning?

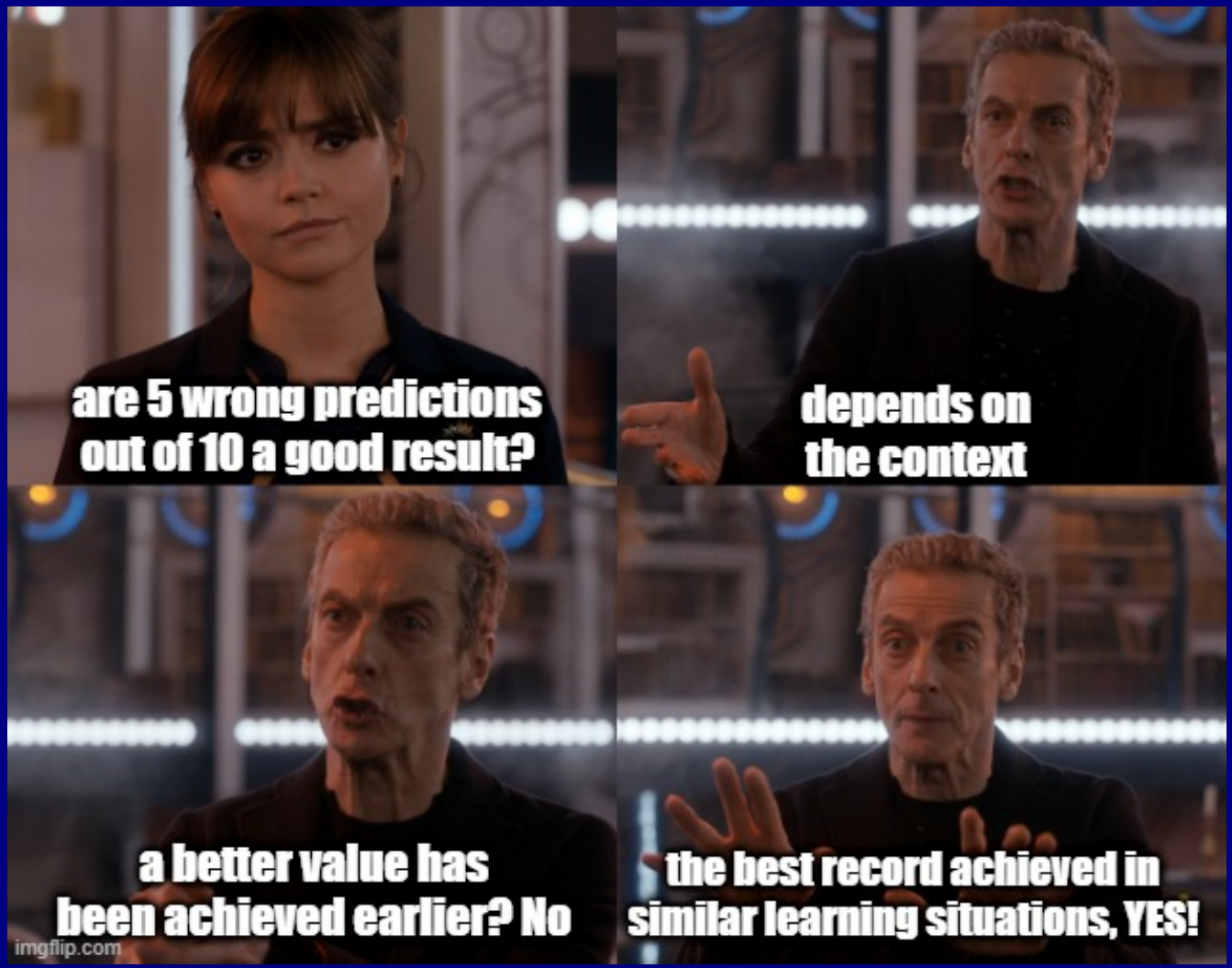


Qualitative approach

- to learn is to acquire expertise with experience accumulated from observations
- we can tell that someone learned when he can make right decisions and judgments
- not every learning is good. sometimes we learn wrong things if we had an ambiguous examples or followed a wrong thinking pattern
- bad examples and bad patterns used for learning are the main reason for a bad learning

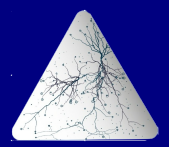


Qualitative approach: Relativity issue...



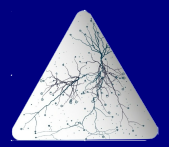
Quantitative approach

- learning is the difference between the # of right decisions / judgments a model has made, and the # those it should had made
- the smaller this difference is, the higher is the learning
- we mathematically define 2 types of learning performance:
  1. **Empirical Loss Error (L_s):** is the ratio of wrong predictions to the # of total predictions made by an algorithm for his given data. It's named empirical because it evaluates the error margin only relatively to the past data "used" in training and testing.
  2. **Generalization Loss Error (L_D):** is a more advanced concept which tries to evaluate if the actual trained model still gives right answers for different forms of dataset. In this function we also define a probability low that rules the attitude of the dataset for the model.



Real-life example: preparing for exams!

Reality	**Mathematical Modeling**
3 students (S_1, S_2, S_3)	3 learning models: A, B, C
1 subject (Graph Theory)	1 knowledge field
courses, labs, and past exams	3 datasets: C, L, E
each student used a resource	each model trained on a dataset
2 different exams	2 different tests

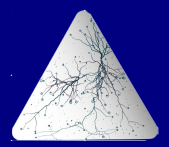


Real-life example: preparing for exams!

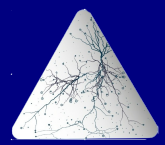
student	**exam 1**	**exam 2**	**model**	**test 1**	**test 2**
S_1	04/20	06/20	A	($L_s =$)80%	70%
S_2	19/20	03/20	B	05%	85%
S_3	14/20	12/20	C	30%	40%

Conclusion

S_1 / A : underfitting ($L_s \sim 1$)
S_2 / B : overfitting ($L_D \sim 1$)
S_3 / C : -relatively- good learning ($L_s \sim 0$ and $L_D \sim 0$)

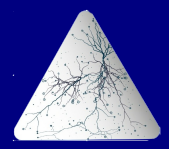


Perfectly Balanced..

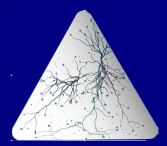


Quantitative approach

- $L_s \sim 0 \iff$ no under-fitting, $L_D \sim 0 \iff$ no over-fitting
- $L_s + L_D \sim 0 \iff$ good learning
- the data used for training should be carefully chosen in order to get a good learning and not bias the results. (no overfitting neither underfitting). Generally, we talk about the **size** and the **features** of the data.
- $\Longrightarrow$ Computational Learning Theory *(TBC)*
- Probabilities, Statistics, Algebra, Analysis, and Optimization.



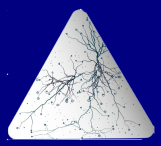
ML obsession



When to use Machine Learning?

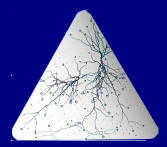
1. some 'underlying pattern' to be learned exists — so 'performance measure' can be improved
2. no programmable (easy) definition — so 'ML' is needed
3. somehow there is data about the pattern — so ML has some 'inputs' to learn from

$\Longrightarrow$ Machine Learning could/should be used. Otherwise..



What are ethics?

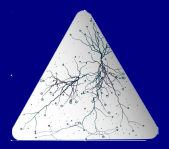
Ethics seeks to answer questions like "what is good or bad", "what is right or what is wrong", or "what is justice, well-being or equality". - *Ethics of AI MOOC*



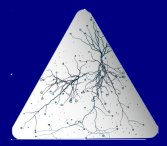
What are AI ethics?

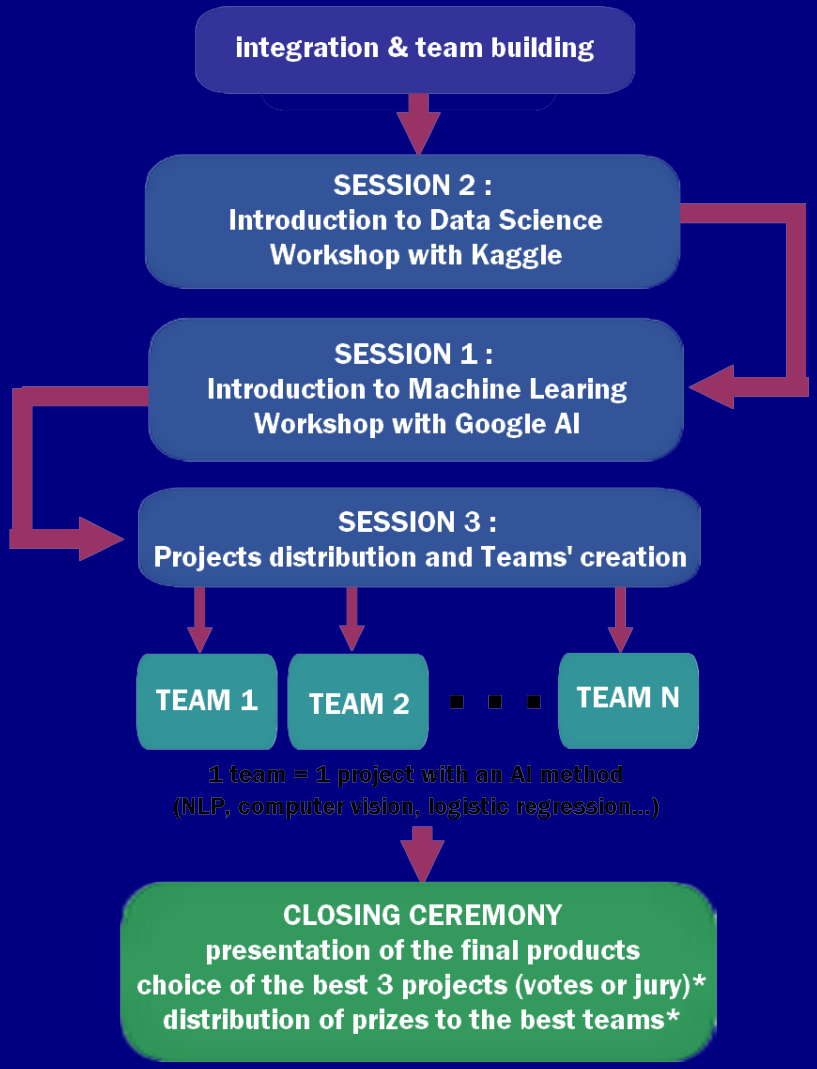
According to a recent study (Jobin et al 2019), AI ethics has quite rapidly converged on a set of five principles:

- non-maleficence
- responsibility or accountability
- transparency and explainability
- justice and fairness
- respect for various human rights (such as privacy and security)



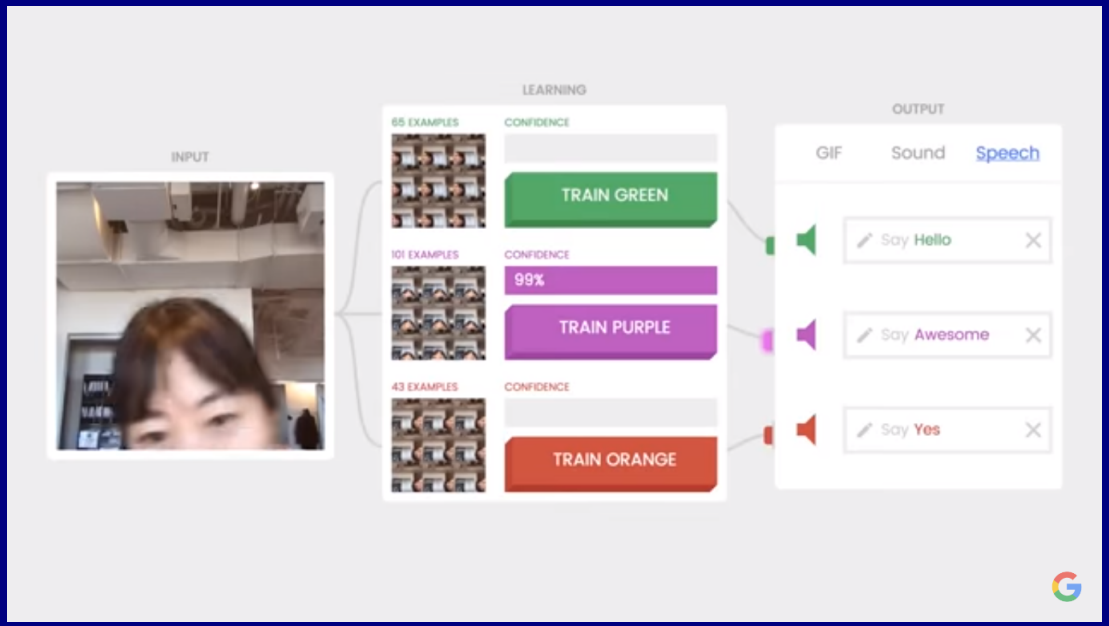
Why we need ethics for AI?

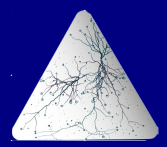



integration & team building
SESSION 2 :
Introduction to Data Science
Workshop with Kaggle
SESSION 1 :
Introduction to Machine Learing
Workshop with Google AI
SESSION 3 :
Projects distribution and Teams' creation
TEAM 1
TEAM 2
TEAM N
1 team = 1 project with an AI method
(NLP, computer vision, logistic regression...)
CLOSING CEREMONY
presentation of the final products
choice of the best 3 projects (votes or jury)*
distribution of prizes to the best teams*

What is Google AI Experiments?

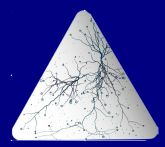
AI Experiments is a showcase for simple experiments that make it easier for anyone to start exploring machine learning, through pictures, drawings, language, music, and more.
link: Teachable Machine



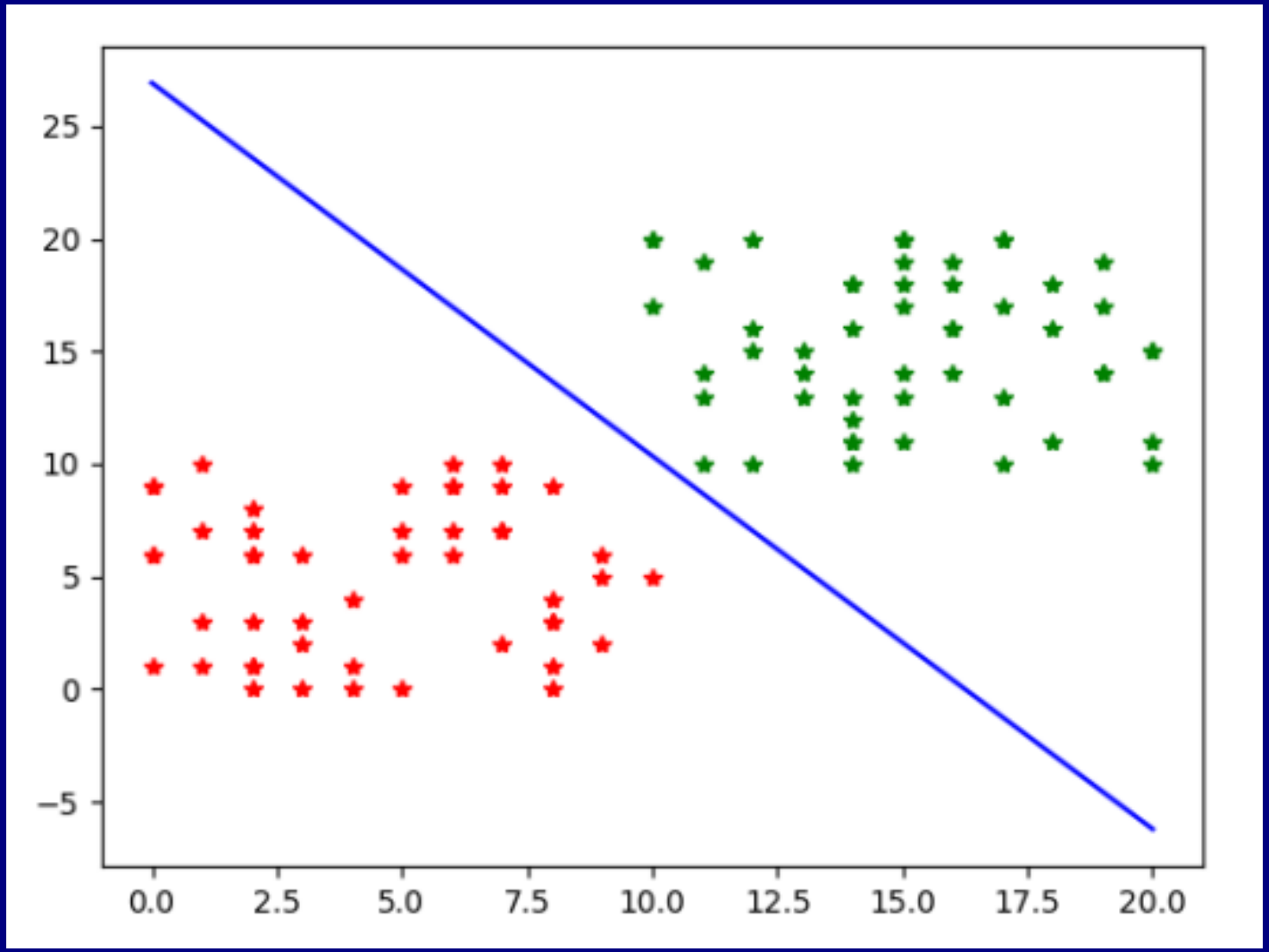


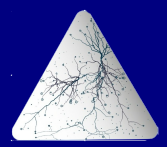
Binary Classification

The objective of this type of ML is to clearly distinguish between 2 categories of data according to multiple features.
Mathematically speaking, for a given set X of vectors $x_i \in \mathbb{R}^n$ and a given function $f : \mathbb{R}^n \longrightarrow \{-1, 1\}$, we try to find an hyperplane $H \subset \mathbb{R}^{n-1}$ that separates the space $\mathbb{R}^n = D$ into 2 sub-spaces D^+ and D^- such that $f(x_i) = 1$ for all $x_i \in D^+$ and $f(x_i) = -1$ for all $x_i \in D^-$.
Since the X's data could be noisy, we should find the optimal hyperplane H that minimize Empirical Loss Error which is in this case, the expectation value of the number of misclassified vectors.



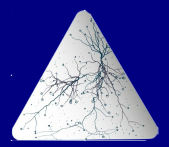
Binary Classification





Perceptron Algorithm

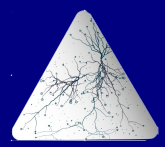
The perceptron was invented in 1943 by McCulloch and Pitts. The first implementation was a machine built in 1958 at the Cornell Aeronautical Laboratory by the psychologist **Frank Rosenblatt**.



Perceptron Algorithm

Algorithm: Perceptron Learning Algorithm

$P \leftarrow$ *inputs with label* 1;
$N \leftarrow$ *inputs with label* 0;
Initialize $\mathbf{w}$ randomly;
while !*convergence* **do**
  Pick random $\mathbf{x} \in P \cup N$;
  if $\mathbf{x} \in P$ *and* $\mathbf{w}.\mathbf{x} < 0$ **then**
    $\mathbf{w} = \mathbf{w} + \mathbf{x}$;
  end
  if $\mathbf{x} \in N$ *and* $\mathbf{w}.\mathbf{x} \geq 0$ **then**
    $\mathbf{w} = \mathbf{w} - \mathbf{x}$;
  end
end
//the algorithm converges when all the inputs are classified correctly



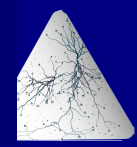
Perceptron Algorithm

TO DO

1. write a function that returns random linearly separable data in 2D, and in 3D
2. write a function for the PLA in any dimension.
   input:
   - X : matrix where the rows are the data records and the columns are the data features.
   - Y : list of labels of the records (label[X[i]] = Y[i], either 1 or -1)

   output: optimal hyperplane coordinates and the number of iteration made to find it
3. write a function that plots the data and the hyperplane separating it for 2D case.

material: `numpy|`, `matplotlib|`, `random|`.



Introduction to Machine Learning and Artificial Intelligence

Hamza BA-MOHAMMED

EAIC Vice President charged of Training

November 10, 2022