# Machine learning: General concepts

Arthur France-Lanord

Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie

PISACMS 2025

École internationale DIADEM 2025

# Outline

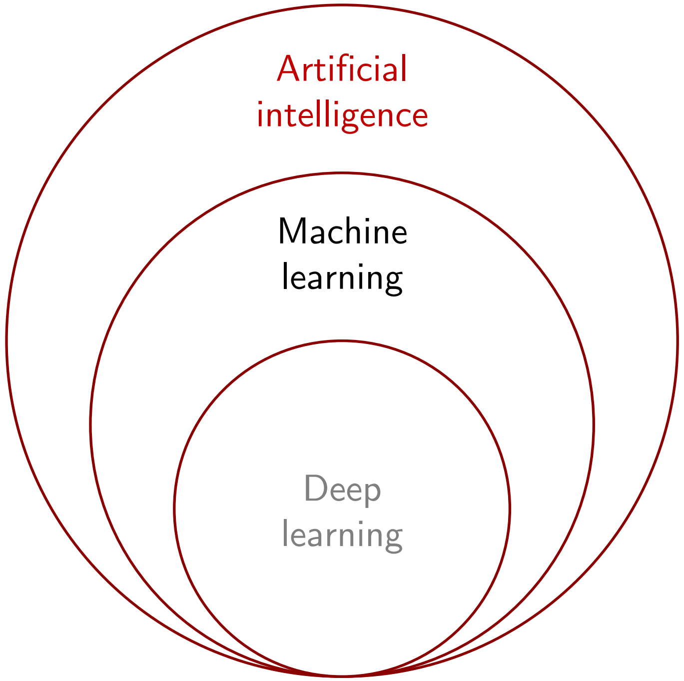
- Context, definitions, keywords: artificial intelligence, machine learning, deep learning
- Different forms of learning
- The heart of ML: data representations and transformations
- Artificial neural networks
- An example from molecular dynamics simulations

# Artifical intelligence – Machine learning – Deep learning

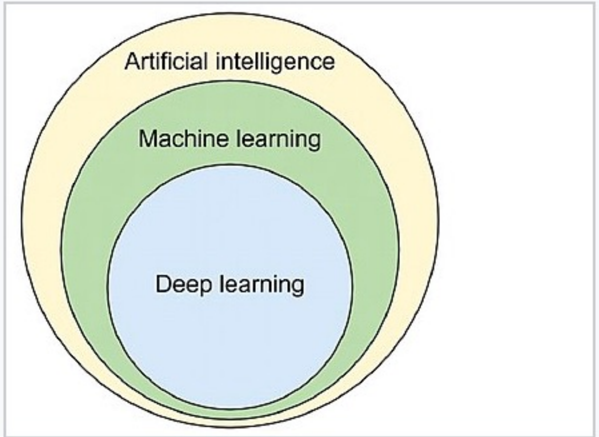
The effort to automate intellectual tasks normally performed by humans

Making decisions or predictions based on data

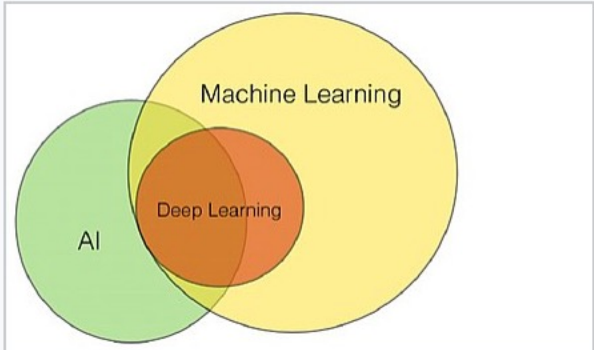
Data is abstracted multiple times



# A word on taxonomy




Machine learning as subfield of AI[18]




Part of machine learning as subfield of AI or part of AI as subfield of machine learning[19]

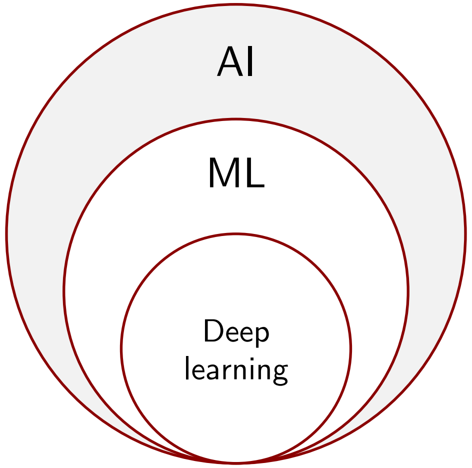
*wikipedia*

Emerged as a field of research in 1956:

*"The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer."*

**John McCarthy**

Proposal for the Dartmouth Summer Research Project on Artificial Intelligence
summer 1956

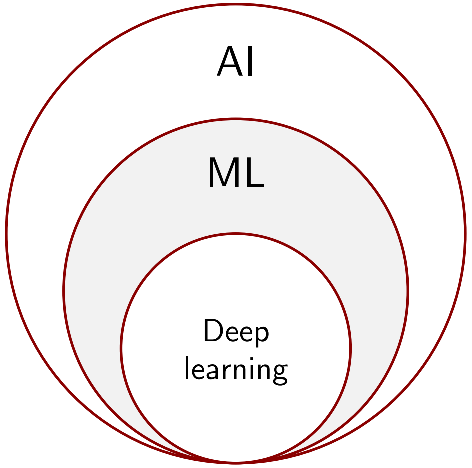


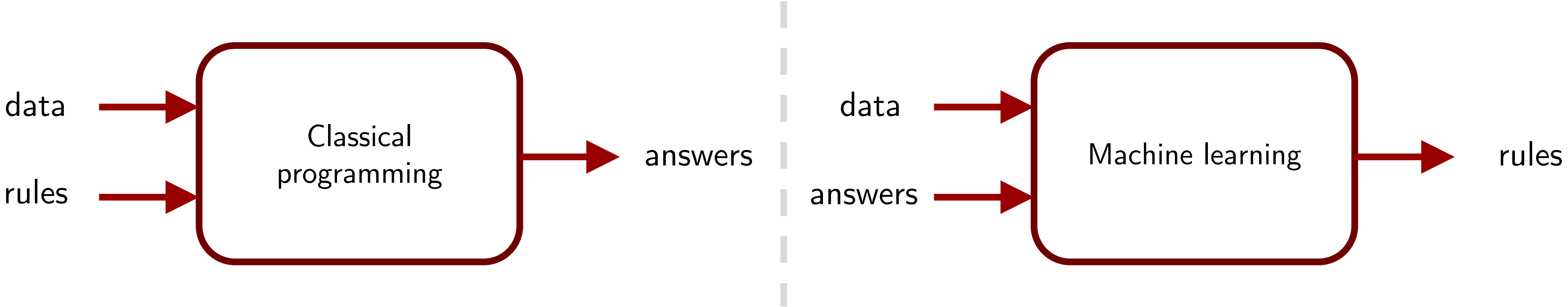

*wikipedia*

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

***Arthur L. Samuel***

“Some studies in machine learning using the game of checkers”
*IBM journal of research and development* 3.3, 1959



# Different forms of algorithms address different forms of problems

**Supervised learning**

Learn patterns from labeled data

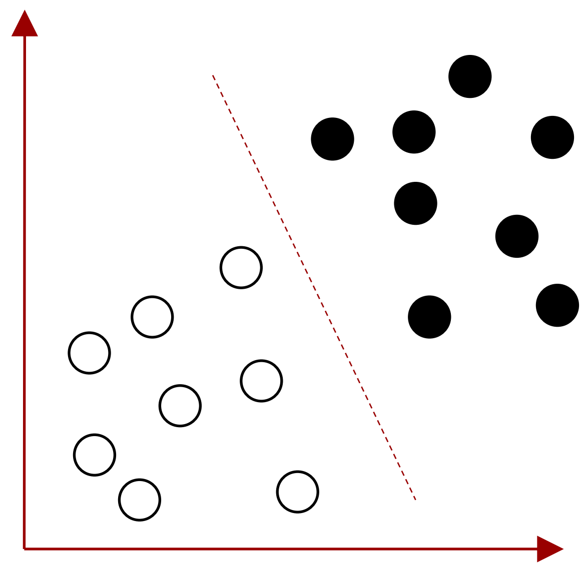
(classification, regression)

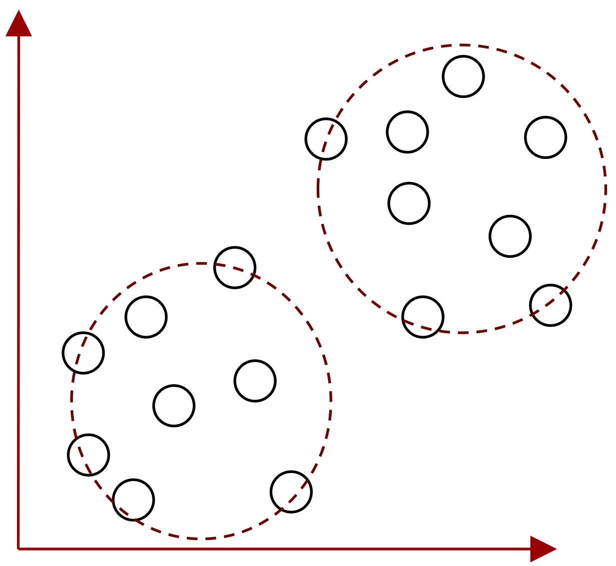
**Unsupervised learning**

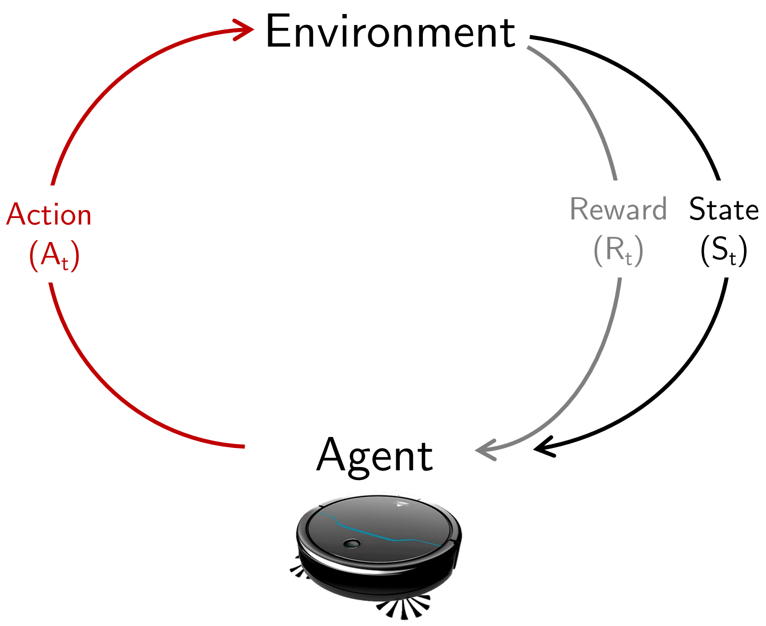
Learn patterns from unlabeled data

**Reinforcement learning**

Learn patterns to make decisions that return the greatest reward based on current and expected future conditions







# Different forms of algorithms: supervised learning

The task of learning a function that maps an input to an output based on example input-output pairs.

input

map

“cat” “cat” “dog”

“dog” “cat” “cat”

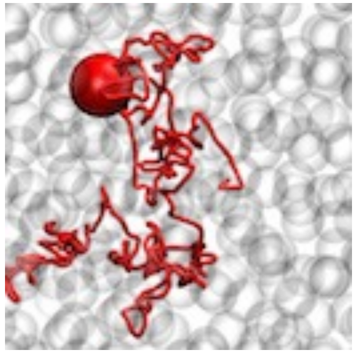
output

“cat” “cat” “dog”

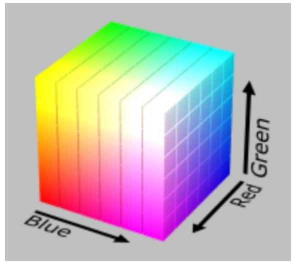
example i-o pairs

# Data representations

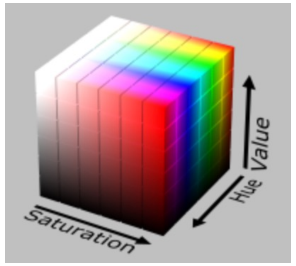
A representation is a different way to look at data.



input (image)




RGB




HSV

easier

harder

easier

"Select all red pixels in the image"

"Make the image less saturated"

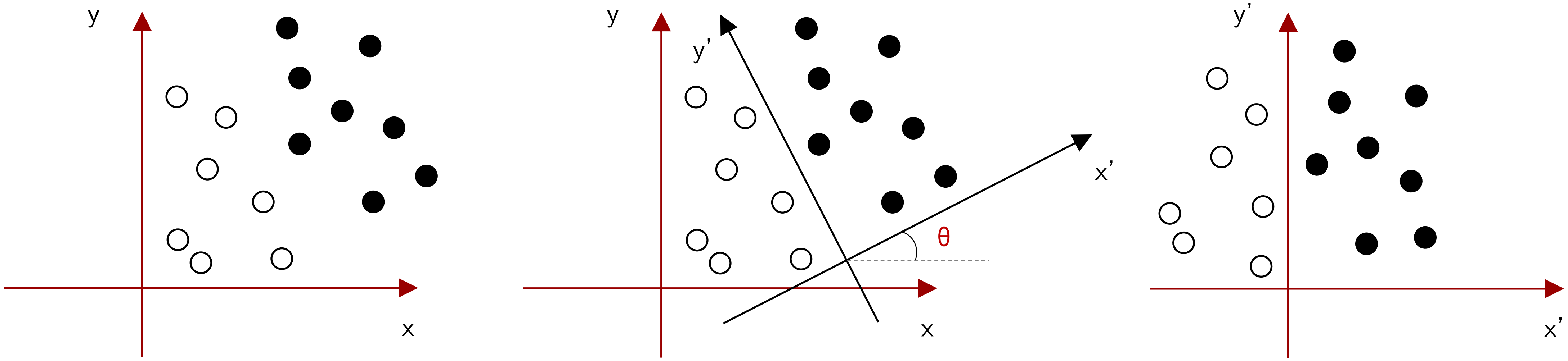
different representations

different problems

# Data representations as transformations

Transformations applied to data lead to new representations.

From their coordinates, classify points according to their color



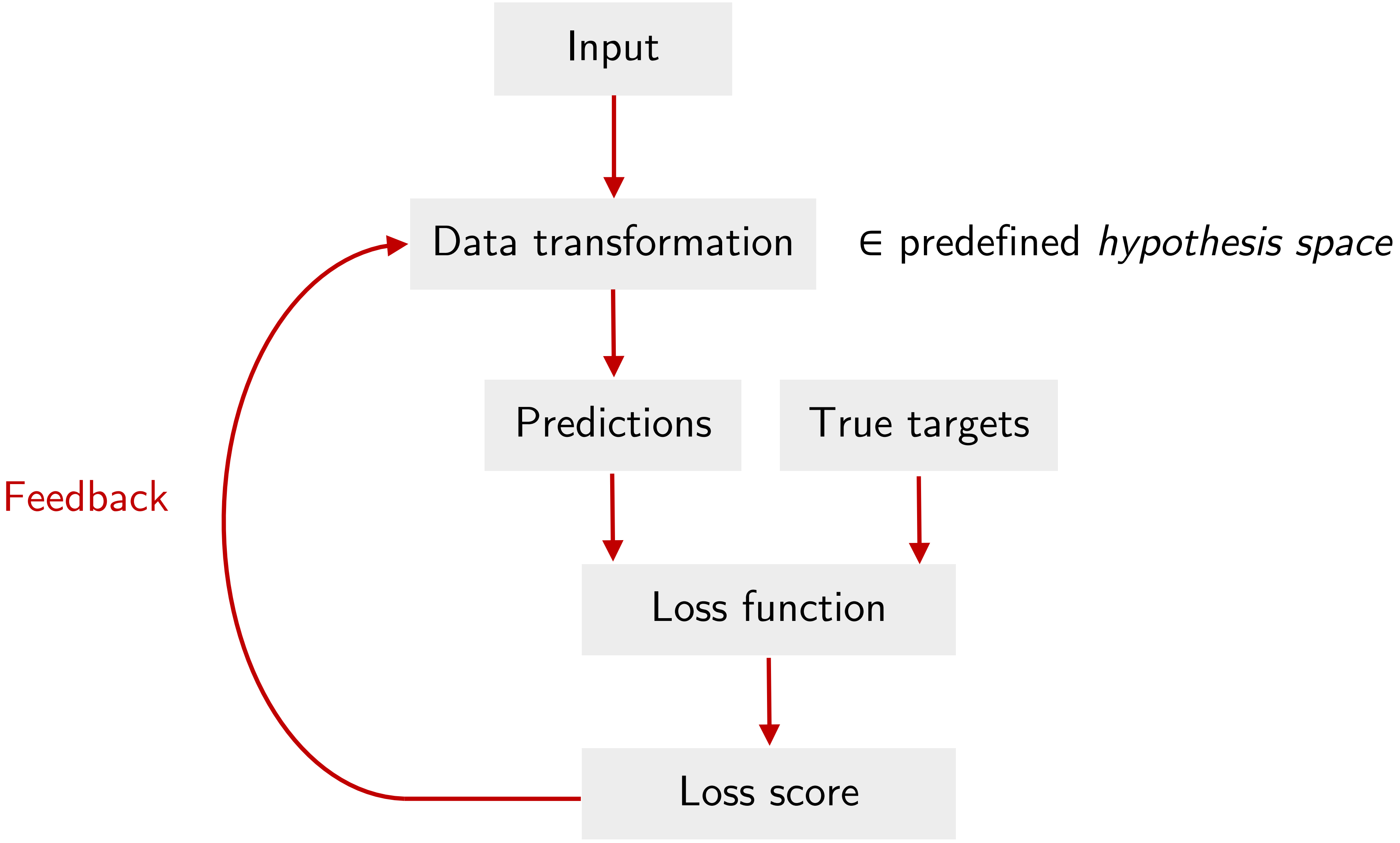

Raw data | Transformation | Better representation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

Black: x' > 0 ; white: x' < 0

Handcrafted representations: feature engineering

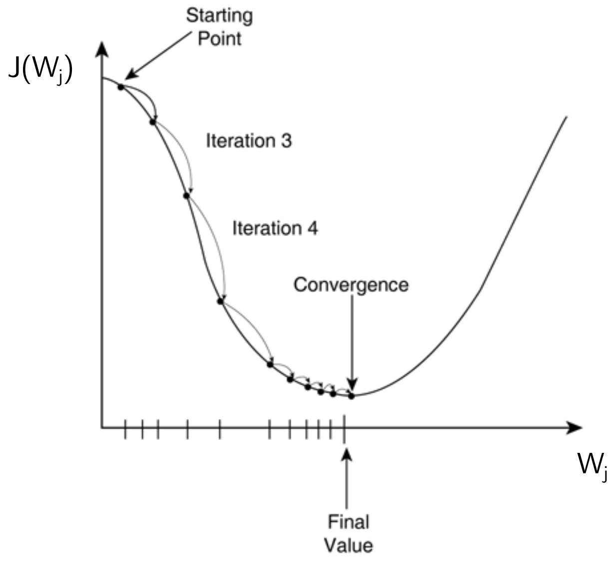
# Where is the learning?




What is learned are meaningful representations of data, through optimization

# Optimizing is learning

The loss function is used to asses the performance of the model.
Weights are updated using an optimizer to minimize loss score.




Different loss functions can be used:

- Mean absolute error
- Mean squared error
- Cross-entropy, *etc*

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i|$$

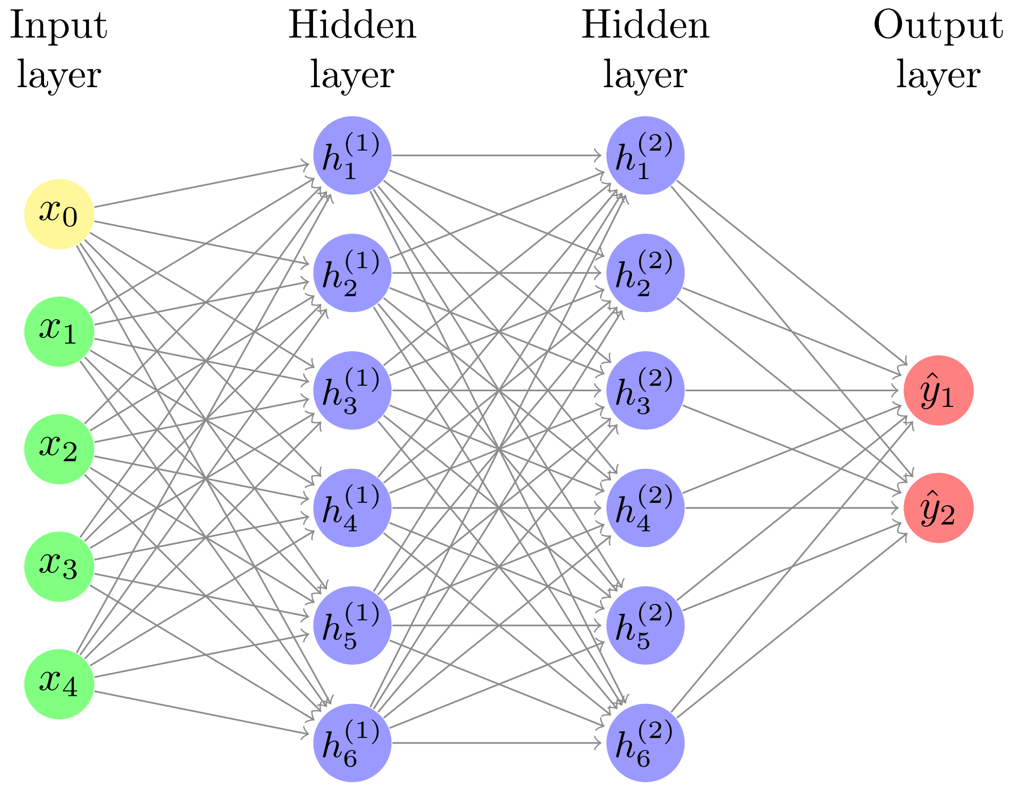
$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

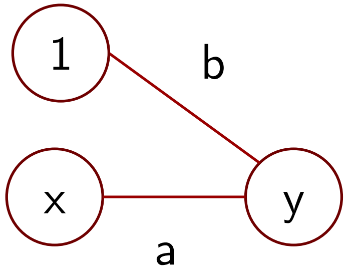
Usually, gradient descent is used to update the weights.

$$W_j \leftarrow W_j - \alpha \frac{\partial}{\partial W_j} J(W)$$

# Artificial neural networks (ANN)

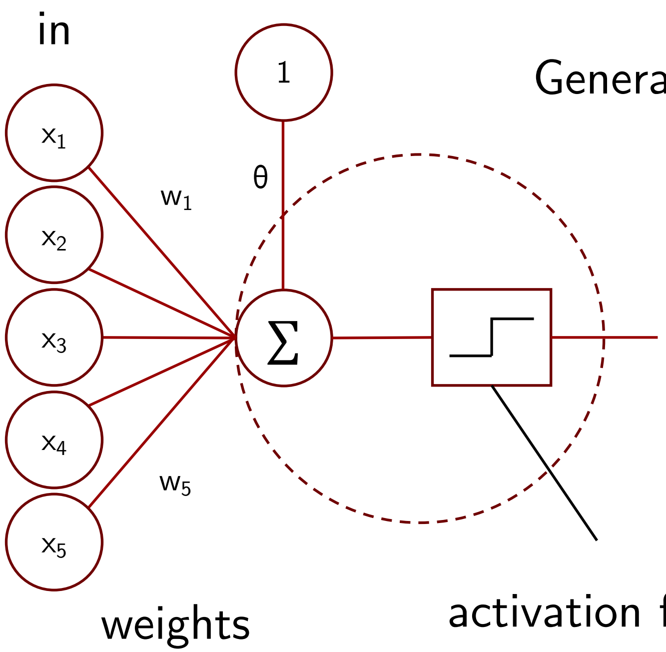
A collection of connected nodes (neurons) in which a signal (data) is transmitted.




Linear network

$y = a \cdot x + b$

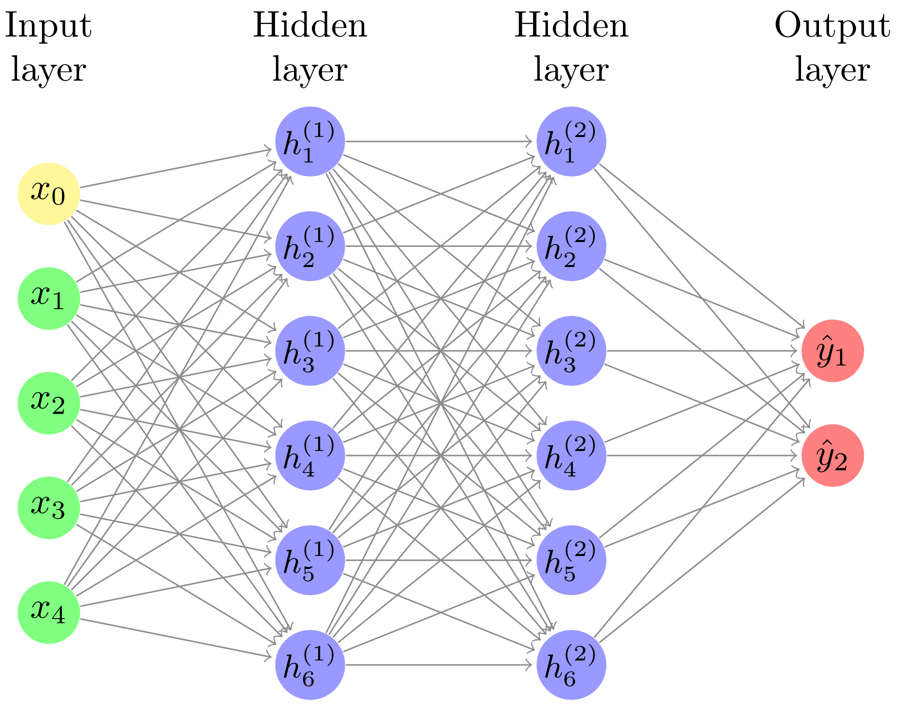


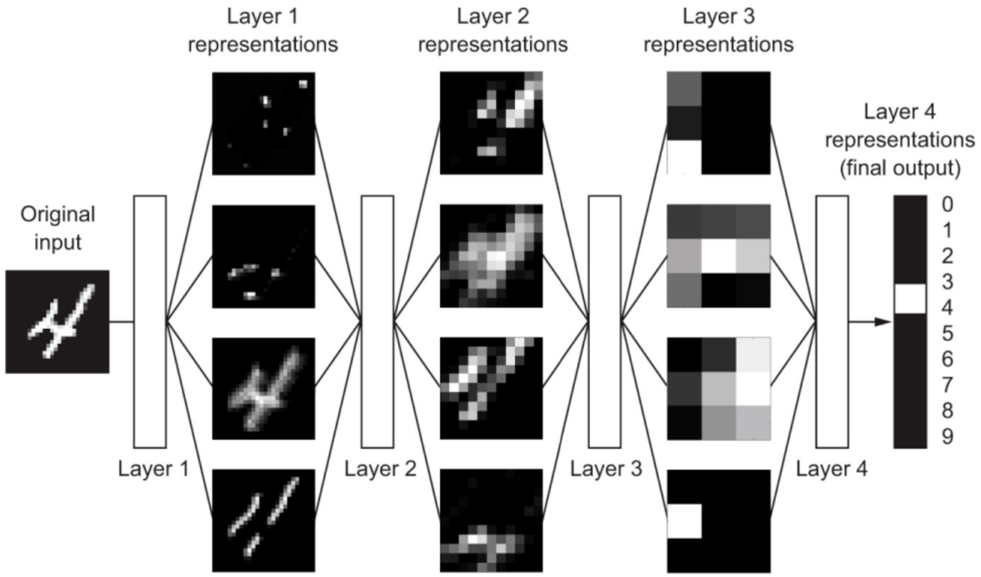

General case

activation function (nonlinear part)

# Multilayer perceptrons (MLP)

- Use at least one hidden layer
- Allow to learn more complex patterns
- Deep learning concerns networks of two or more hidden layers

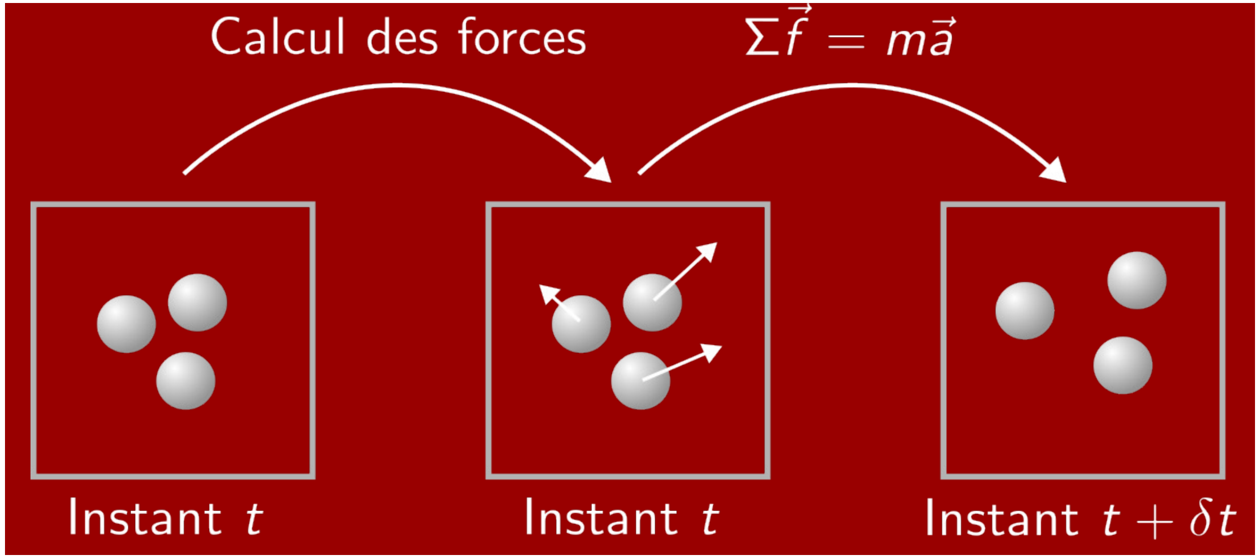


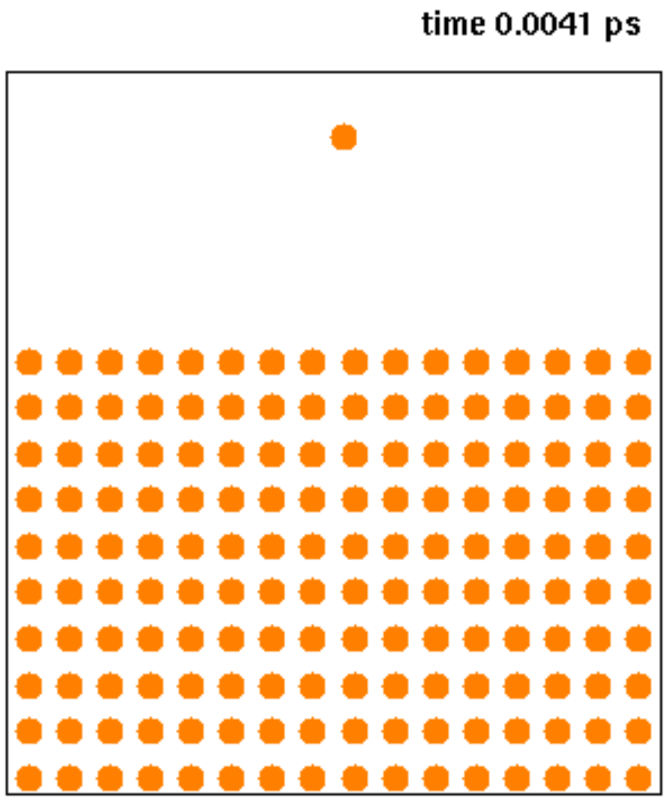




*François Chollet*

# Molecular dynamics and machine-learned interatomic potentials

# Molecular dynamics




*wikipedia*

Output: trajectory (basically a movie of the system)

$$\mathbf{M}\ddot{\mathbf{X}} = F(\mathbf{X}),$$
$$F(\mathbf{X}) = -\nabla U(\mathbf{X})$$

?

# Machine-learned interatomic potentials (MLIPs)

Computational cost

Classical potentials | *ab initio*

$$E = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

$$E_0 = \langle \Psi[n_0] | \hat{\mathcal{H}} | \Psi[n_0] \rangle$$

Accuracy

MLIPs are an attempt to fill the zone in between classical potentials and *ab initio* methods:

- Good accuracy, as close to *ab initio* as possible
- Reasonable computational cost, especially low complexity

# What is a machine-learned interatomic potential?

Interatomic potential: $$V = f(\mathbf{X})$$

A typical form: $$\{\mathbf{X}_i, \mathbf{X}_j\} \to r, \quad r = |\mathbf{X}_i - \mathbf{X}_j|$$

$$V(r) = 4\varepsilon \left[ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right]$$

Machine learning variant (MLIP):

$$\mathbf{X} \to \mathbf{D} \qquad V = f(\mathbf{D})$$

Data representation — Mapping function

ACSF, SOAP, ... — NN, Gaussian process, ...

Now let's try it out!

Computer lab session with **João Paulo Mendonça**