Introduction to Machine Learning

Francesco Pugliese, PhD

neural1977@gmail.com

Che cosa è il Machine Learning ?

- ✓ Machine learning è un ramo dell'Informatica che riguarda la costruzione di algoritmi, che si fondano su una collezione di dati riguardanti un determinate fenomeno.

- ✓ Il termine Machine Learning fu coniato da Arthur Samuel in 1959, un Americano pioniere nel campo dei videogiochi e dell'Intelligenza Artificiale che affermava: "**Il Machine Learning fornisce ai computer la capacità di apprendere senza essere programmatic esplicitamente**".

- ✓ Machine learning è uno strumento per la trasformazione dell'**Informazione in Conoscenza**. Gli schemi nascosti (pattern) e la conoscenza concernente un determinate problema possono essere utilizzati per predire eventi future ed eseguire tutti i tipi di decision making complessi.

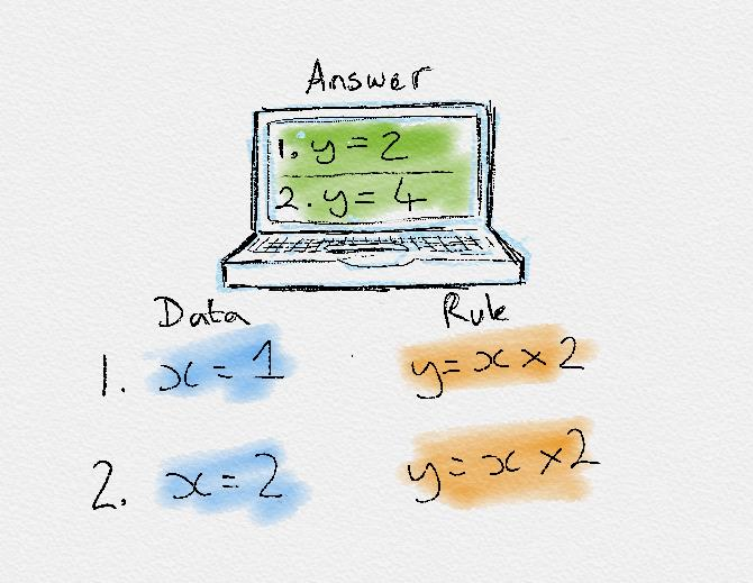
- ✓ Il Machine learning può anche essere definite come il processo per risolvere un problema pratico attraverso:
  - ✓ L'acquisizione di un dataset.
  - ✓ Costruendo algoritmicamente un modello statistic basato su quel dataset.

  Quel modello statistico lo si assume essere in qualche modo capace di risolvere il problema pratico.

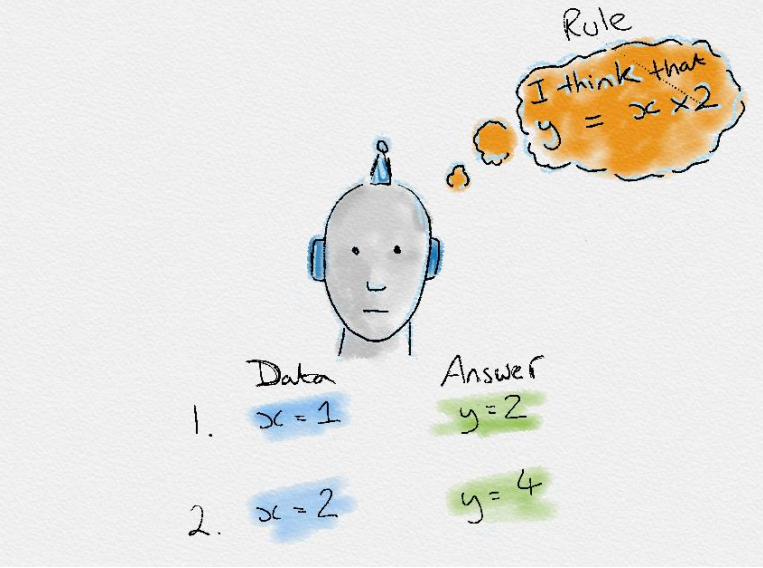
Programmazione Tradizionale vs Machine Learning

- ✓ Tradizionalmente, l'ingegneria del software combina le regole create dall'uomo con i dato al fine di creare delle risposte ad un problema specific. Invece il Machine Learning usa i dati e le risposte per scoprire le regole sottostanti ad un determinate problema.

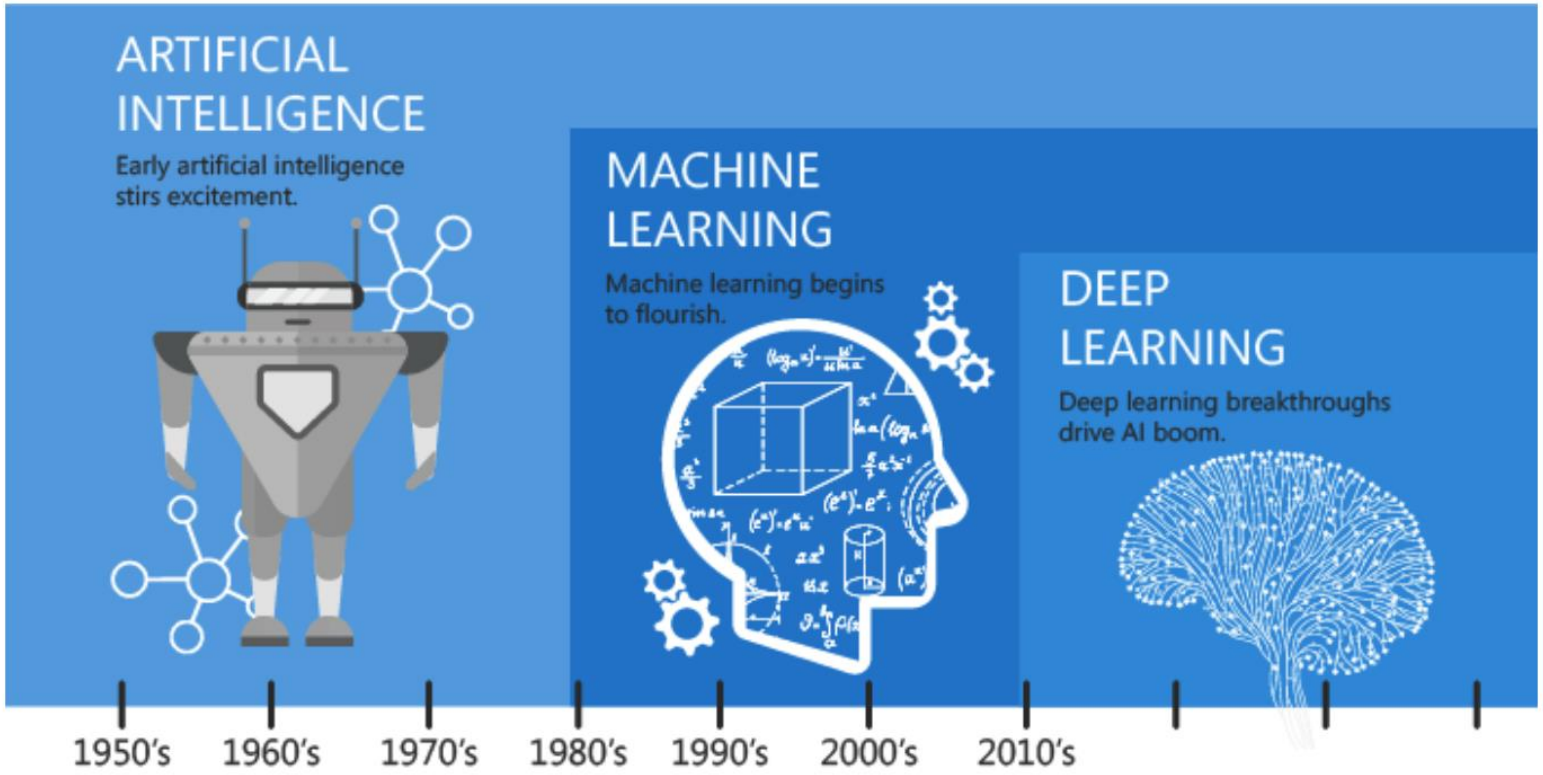
- ✓ Per apprendere le regole che governano un fenomeno, le machine devono attraversare un processo di apprendimenti, provando differenti regole e cercando di capire come esse performano. Ecco perchè esso viene conosciuto come Machine Learning.



VS



Una breve storia del Machine Learning



Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

Image: Linked In | Machine Learning vs Deep learning

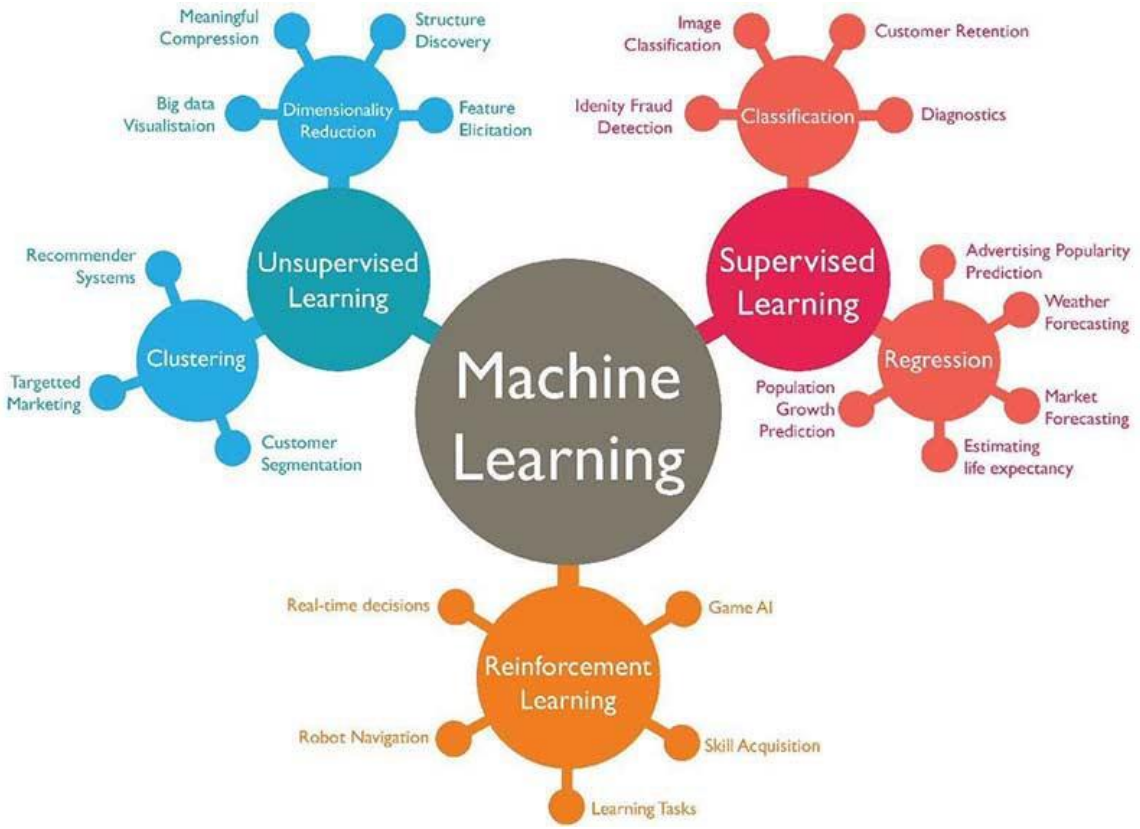
Tipi di Machine Learning

- ✓ Esistono 4 tipi di Machine Learning ampoa,emte riconosciuti:
  - ✓ Supervised learning
  - ✓ Unsupervised learning
  - ✓ Semi-supervised learning
  - ✓ Reinforcement learning
- ✓ Ciascuna forma di Machine Learning presenta diversi approcci, ma tutti seguono lo stesso processo e la stessa teoria sottostante.
- ✓ Il **No Free Lunch theorem** è famoso nel Machine Learning. Questo teorema afferma che non esiste un singolo algoritmo che lavora bene per tutti i tipi di task. Ciascun task che cerchi di risolvere ha le sue proprie idiosincrasie, ovvero le proprie specificità ed avversità a determinati modelli del Machine Learning. Quindi, ci sono molti algoritmi e approcci che si adattano a ciascun specificità e stranezza di un singolo problema.

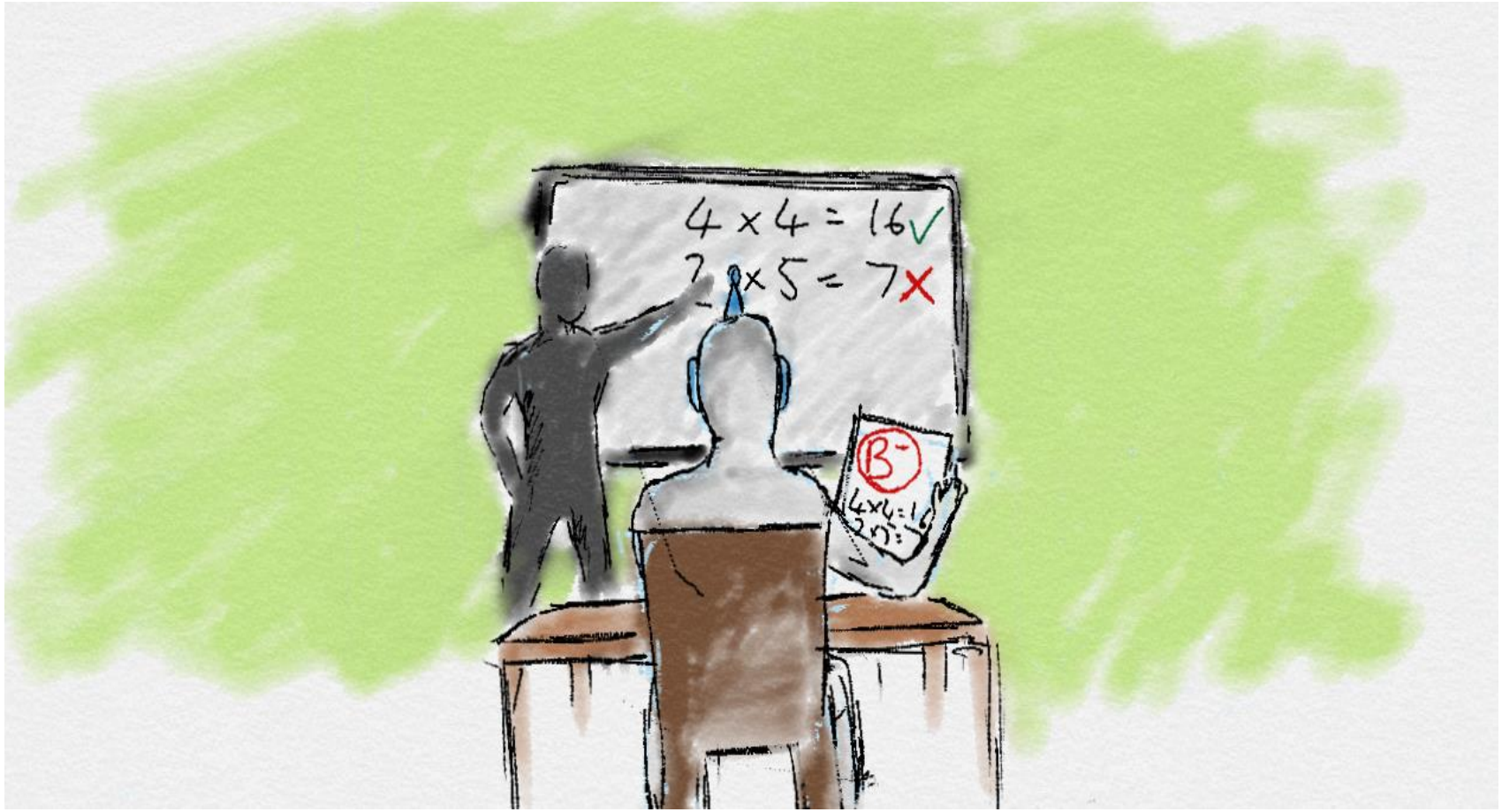
No Free Lunch Theorem

- ✓ Il **No Free Lunch Theorem (NFL o NFLT)** è spesso utilizzato nel campo dell'ottimizzazione e del Machine Learning, ed afferma che tutti gli algoritmi di ottimizzazione performano ugualmente bene quando è calcolata la media delle loro performance su tutti i possibili problemi che possono esistere.
- ✓ Questo implica che non esiste il miglior algoritmo di ottimizzazione i assoluto. E quindi che non eiste il miglior algoritmo di Machine Learning in assoluto per i problem di modellazione predittiva come la classificazione o la regression.
- ✓ Dunque su tutto lo spazio di tutti i possibili problem, ogni tecnica di ottimizzazione performerà mediamente bene come ogni altra tecnica (inclusa la Random Search), e quindi la prima implicazione è che sia per problemi di ottimizzazione statica e dipendente dal tempo la performance media di ogni possibile coppia di algoritmi attraverso tutti i problemi è esattamente identica.

Tipi di Machine Learning



Supervised learning



Supervised learning

- ✓ In supervised learning, the goal is to learn the **mapping** (the rules) between a set of inputs and outputs.
- ✓ For example, the inputs could be the **weather forecast**, and the outputs would be the **visitors** to the beach.
- ✓ The goal in supervised learning would be to learn the mapping that describes the relationship between temperature and number of beach visitors.
- ✓ Example labelled data is provided of past input and output pairs during the learning process to teach the model how it should behave, hence, "supervised" learning.
- ✓ For the beach example, new inputs can then be fed in of forecast temperature and the Machine learning algorithm will then output a future prediction for the number of visitors.

Supervised learning

- ✓ Being able to adapt to new inputs and make predictions is the crucial **generalisation** part of machine learning. In training, we want to **maximise generalisation**, so the supervised model defines the real "general" underlying relationship.

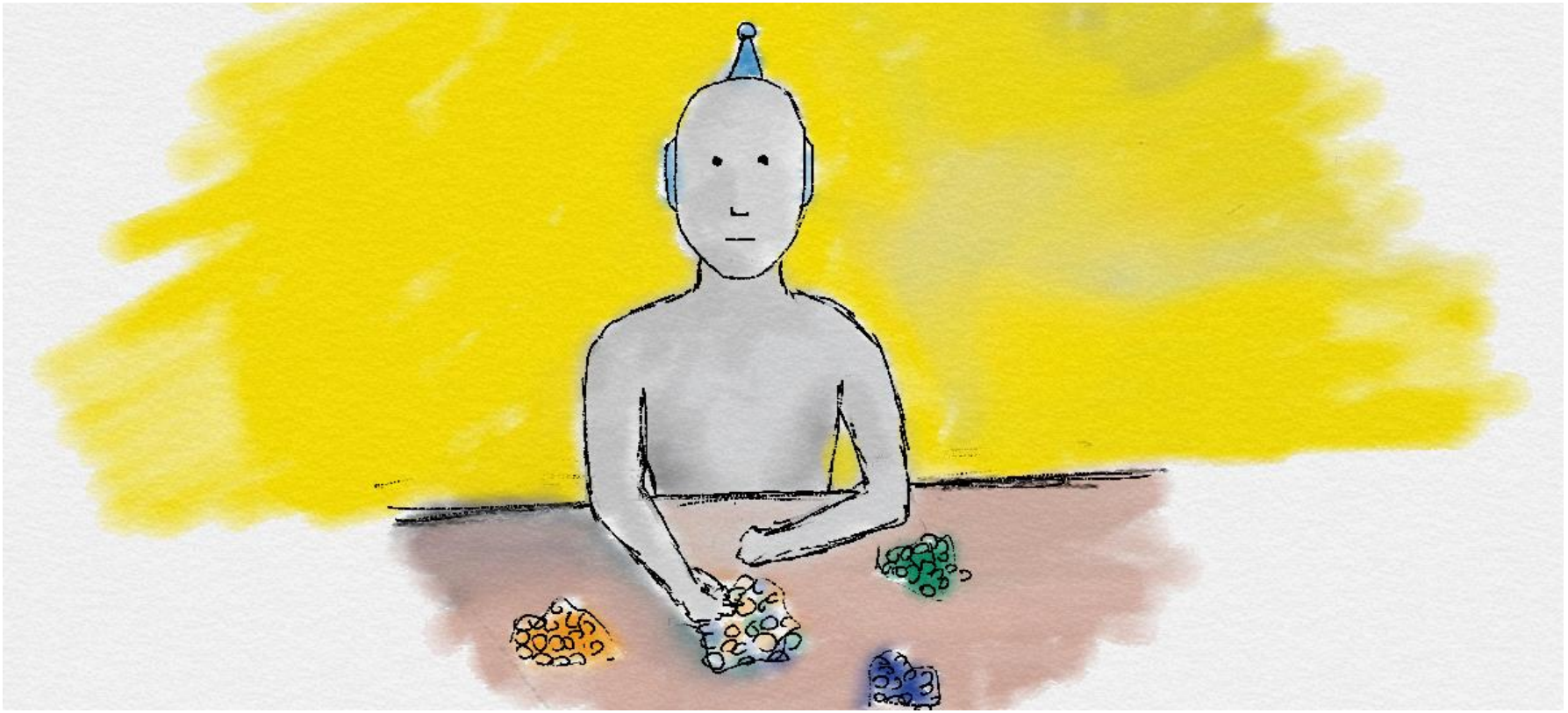
- ✓ If the model is over-trained, we cause over-fitting to the examples used and the model would be unable to adapt to new, previously unseen inputs.

- ✓ A side effect to be aware of in supervised learning that the supervision we provide introduces **bias** to the learning. The model can only be imitating exactly what it was shown, so it is very important to show it **reliable**, **unbiased examples**.

- ✓ Also, supervised learning usually requires a lot of data before it learns.

- ✓ Obtaining enough **reliably labelled data** is often the hardest and most expensive part of using supervised learning. (Hence why **data has been called the new oil**!)

Unsupervised learning



Unsupervised learning

- ✓ In unsupervised learning, only input data is provided in the examples. There are **no labelled** example outputs to aim for. But it may be surprising to know that it is still possible to find many interesting and complex patterns hidden within data without any labels.

- ✓ An example of unsupervised learning in real life would be **sorting different colour coins into separate piles**. Nobody taught you how to separate them, but by just looking at their features such as colour, you can see which colour coins are associated and cluster them into their correct groups.

- ✓ Unsupervised learning can be **harder** than supervised learning, as the removal of supervision means the problem has become **less defined**.

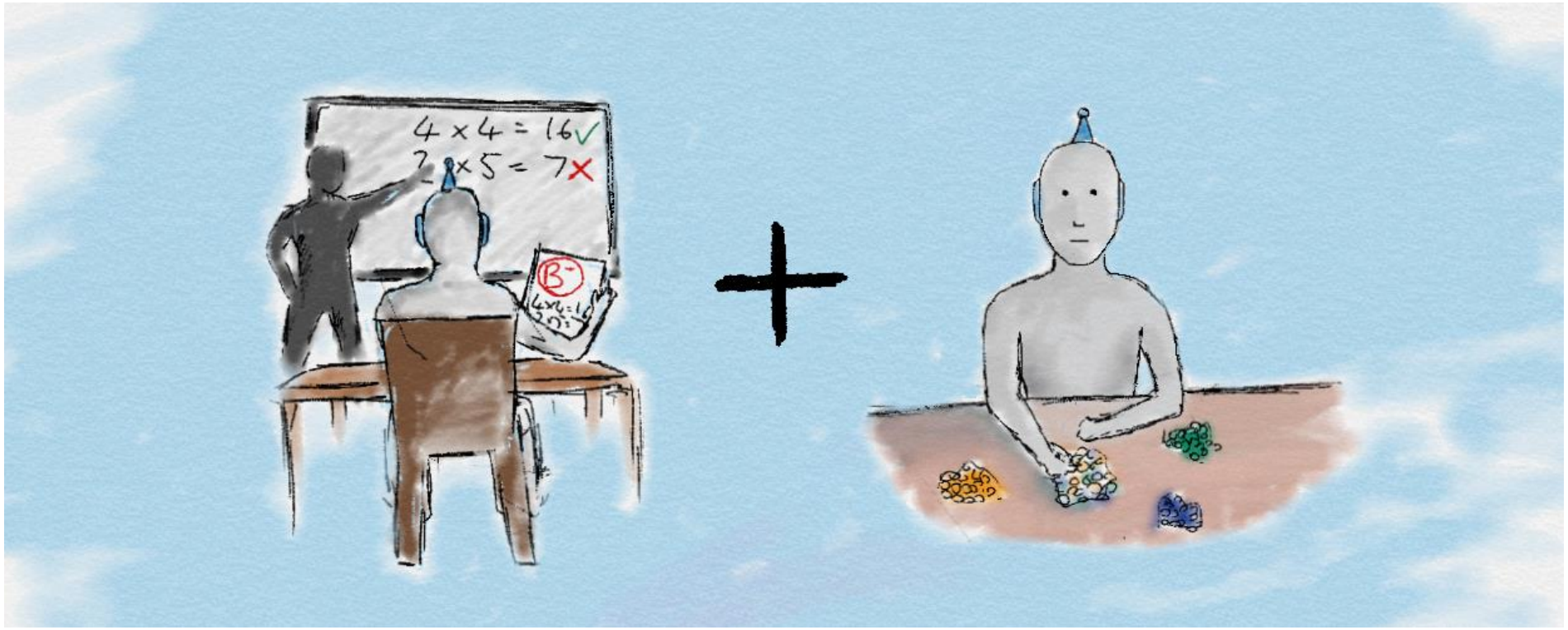
- ✓ You start from a **clean slate** with less bias and may even find a new, better way solve a problem. Therefore, this is why unsupervised learning is also known as **knowledge discovery**. Unsupervised learning is very useful when **conducting** exploratory data analysis.

Techniques of Unsupervised learning

- ✓ Some types of unsupervised learning are
  - ✓ Clustering
  - ✓ Density estimation
  - ✓ Dimensionality reduction
  - ✓ Latent variable models
  - ✓ Anomaly detection

- ✓ More complex unsupervised techniques involve neural networks like Auto-encoders and Deep Belief Networks

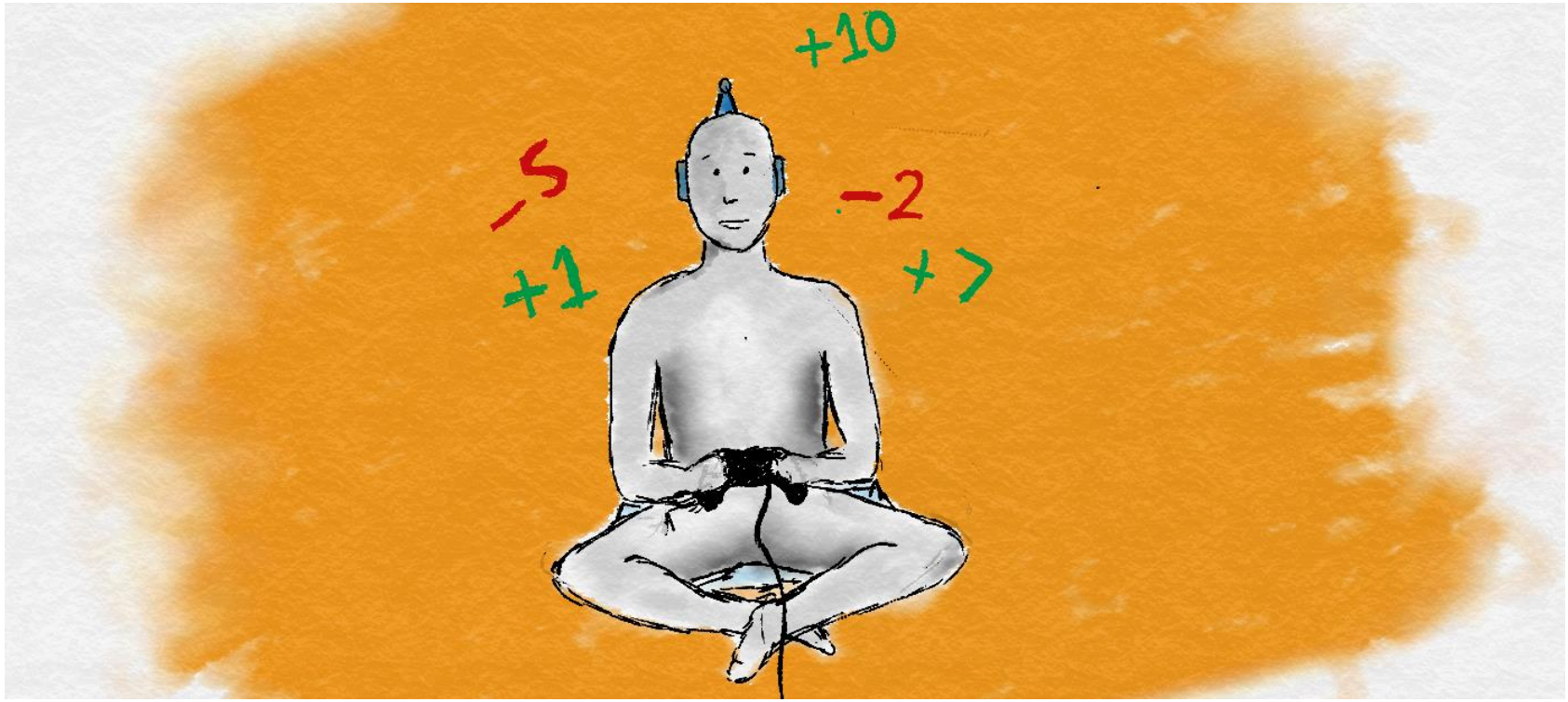
Semi-supervised learning



Semi-supervised learning

- ✓ Semi-supervised learning is a **mix** between supervised and unsupervised approaches.
- ✓ The learning process isn't closely supervised with example outputs for every single input, but we also don’t let the algorithm do its own thing and provide no form of feedback.
- ✓ Semi-supervised learning takes the middle road. By being able to mix together a small amount of labelled data with a much larger unlabelled dataset it **reduces the burden of having enough labelled data**. Therefore, it opens up many more problems to be solved with machine learning.
- ✓ **Example**: A perfect example is in **medical scans**, such as breast cancer scans. A trained expert is needed to label these which is time consuming and very expensive. Instead, an expert can label just a small set of breast cancer scans, and the semi-supervised algorithm would be able to leverage this small subset and apply it to a larger set of scans.

Reinforcement learning



Reinforcement learning

- ✓ Reinforcement learning doesn't use labels as such, and instead **uses rewards to learn**.
- ✓ If you're familiar with psychology, you'll have heard of reinforcement learning. If not, you'll already know the concept from how we learn in everyday life.
- ✓ In this approach, **occasional positive and negative feedback is used to reinforce behaviours**.
- ✓ Think of it like training a **dog**, **good behaviours are rewarded with a treat** and become more common. **Bad behaviours are punished and become less common**. This reward-motivated behaviour is key in reinforcement learning.
- ✓ This is very similar to how we as humans also learn. Throughout our lives, we receive positive and negative signals and constantly learn from them. The chemicals in our brain are one of many ways we get these signals. When something good happens, the neurons in our brains provide a hit of positive neurotransmitters such as dopamine which makes us feel good and we become more likely to repeat that specific action.

Reinforcement learning

Reinforcement learning

- ✓ We don't need constant supervision to learn like in supervised learning. By only giving the occasional reinforcement signals, we still learn very effectively.

- ✓ One of the most exciting parts of Reinforcement Learning is that is a **first step away from training on static datasets**, and instead of being able to use dynamic, noisy data-rich environments. This brings Machine Learning closer to a learning style used by humans. The world is simply our noisy, complex data-rich environment.

- ✓ Games are very popular in Reinforcement Learning research. They provide ideal data-rich environments. A Reinforcement Learning algorithm just aims to maximise its rewards by playing the game over and over again.

- ✓ If you can frame a problem with a frequent 'score' as a reward, it is likely to be suited to Reinforcement Learning.

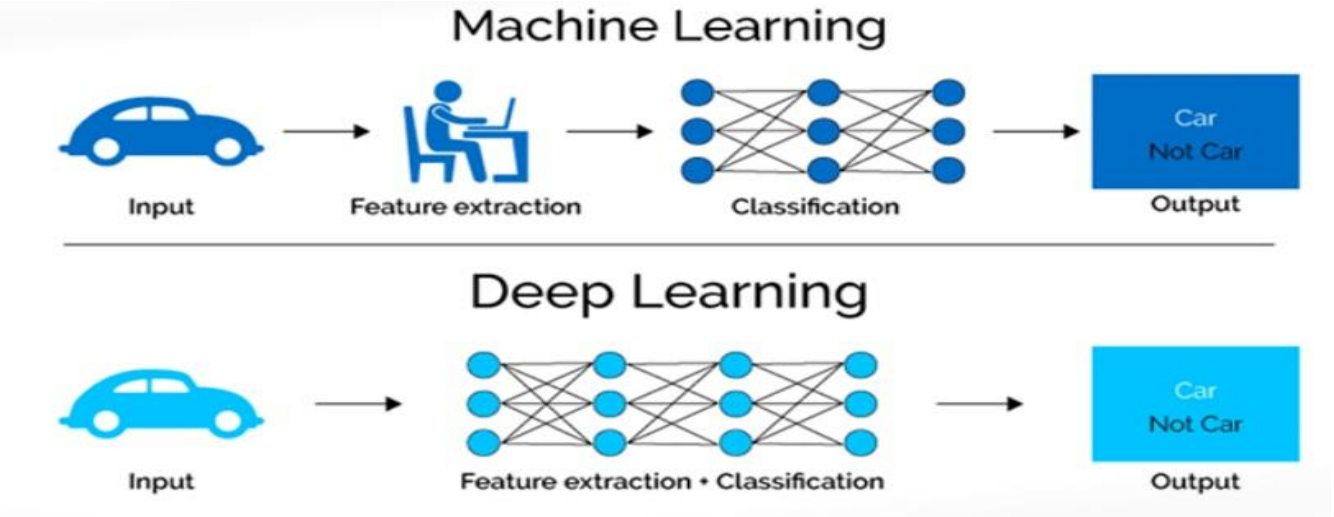
Shallow vs Deep learning

- ✓ A shallow learning algorithm learns the parameters of the model **directly from the features of the training examples**.
- ✓ Most supervised learning algorithms are shallow.
- ✓ The notorious exceptions are **neural network learning algorithms**, specifically those that build neural networks with **more than one layer** between input and output.
- ✓ Such neural networks are called deep neural networks.
- ✓ In deep neural network learning (or, simply, deep learning), contrary to shallow learning, most model parameters are learned not directly from the features of the training examples, but from the **outputs of the preceding layers**.
- ✓ More details about this when we talk about deep learning…

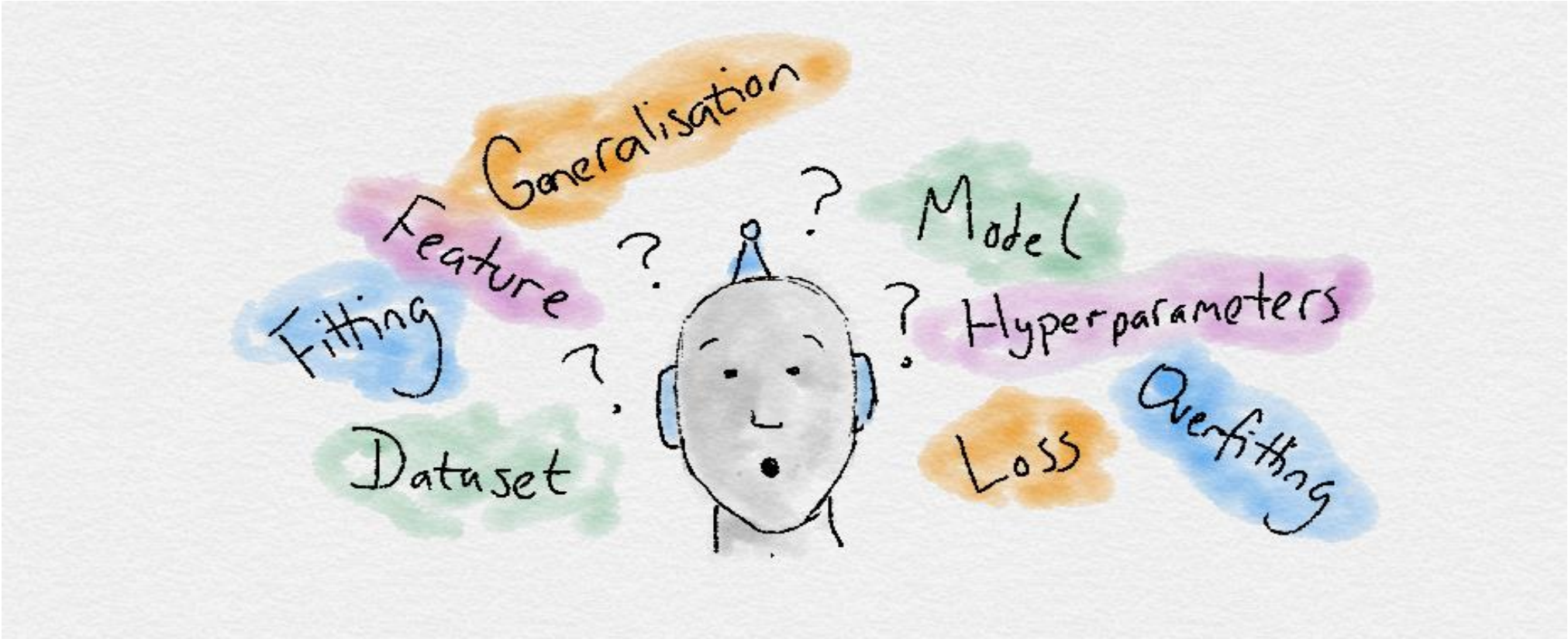
Deep Learning is Machine Learning

Deep Learning refers to algorithms that automatically ‘model’ high-level abstractions in data

i. here ‘model’ means: define, find, recognize and exploit
ii. here ‘automatically’ means: directly from data, without hinging upon handcrafted, task-specific features.



Machine Learning terminologies



Machine Learning terminologies

- ✓ **Dataset**: A set of data examples, that contain features important to solving the problem.

- ✓ **Features**:
  - ✓ Important pieces of data that help us understand a problem.
  - ✓ These are fed in to a Machine Learning algorithm to help it learn.

- ✓ **Model**:
  - ✓ The representation (internal model) of a phenomenon that a Machine Learning algorithm has learnt. It learns this from the data it is shown during training.
  - ✓ The model is the output you get after training an algorithm.
  - ✓ For example, a decision tree algorithm would be trained and produce a decision tree model.

Parameter vs Hyperparameters

- ✓ A hyperparameter is a **property** of a learning algorithm, usually (but not always) having a numerical value. That value **influences** the way the algorithm works.
- ✓ Hyperparameters aren't learned by the algorithm itself from data. They have to be **set by the data analyst before running the algorithm**. More details when we discuss how we tune Machine learning models in the next sections.
- ✓ Parameters are **variables that define the model** learned by the learning algorithm.
- ✓ Parameters are **directly modified** by the learning algorithm based on the training data.
- ✓ The goal of learning is to find such values of parameters that make the model **optimal** in a certain sense.

References

Nolfi, S., Floreano, D., & Floreano, D. D. (2000). *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines*. MIT press.

Pugliese, F., Acerbi, A., & Marocco, D. (2015). Emergence of leadership in a group of autonomous robots. *PloS one*, *10*(9), e0137234.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, *2*(1), 1-127.

Francesco **Pugliese**