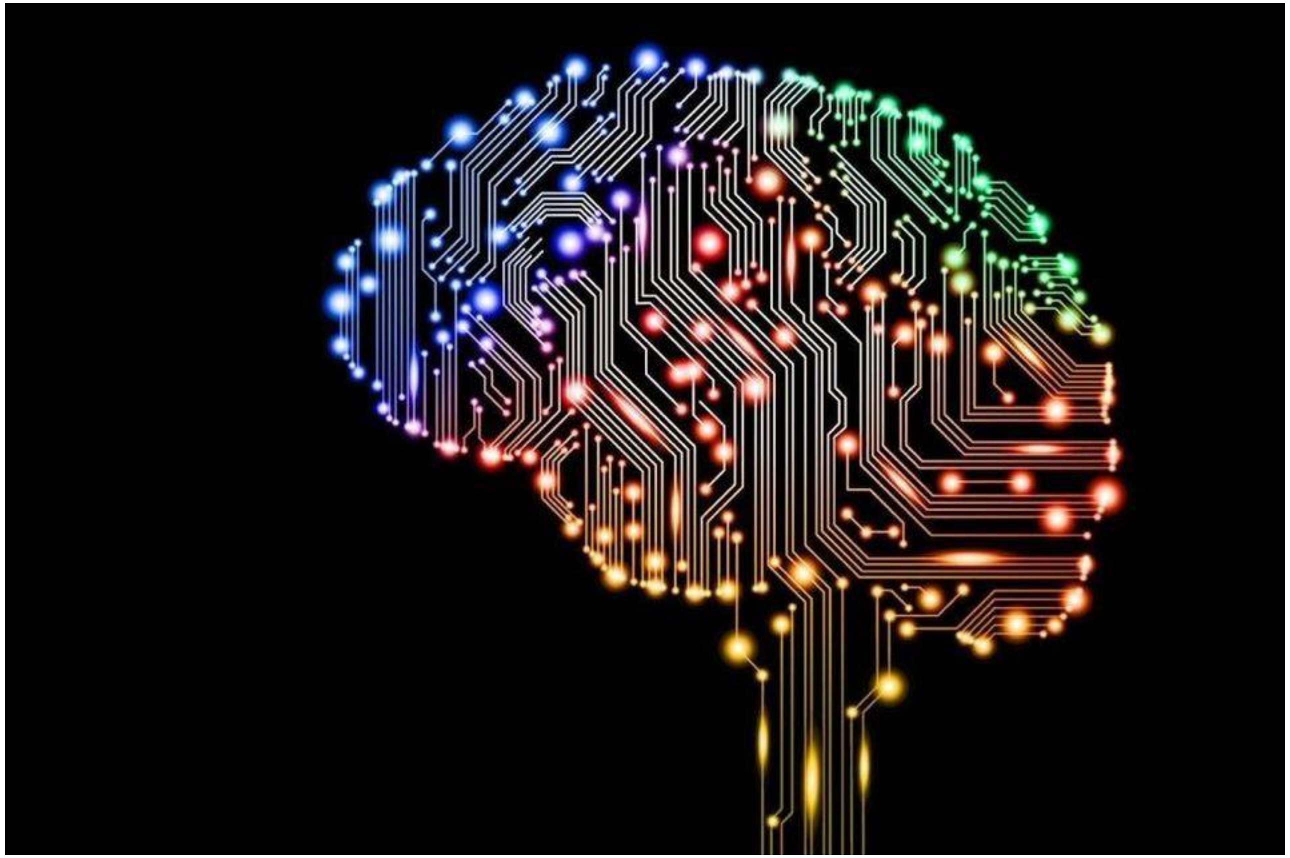
AI - introduction

Dr. Stefaan Haspeslagh



When no data is available...

vives University of Applied Sciences

Who am I?

- Research manager @ Vives University of applied sciences and KU Leuven
  - Topics: AI – machine learning – deep learning
  - Applications in industry, biomedical sciences, logistics, …

- Background in metaheuristic, exact and hybrid algorithms to tackle combinatorial optimisation problems (planning, scheduling, cutting, packing, vehicle routing, group composition, …)

- Lecturer in AI: optimisation – ML – DL

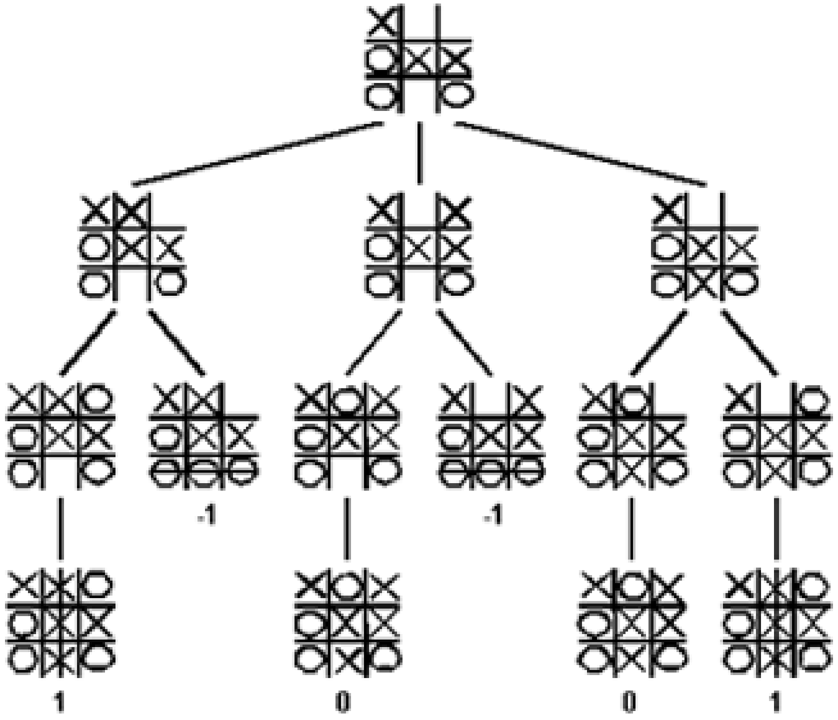
Artificial Intelligence: introduction and examples

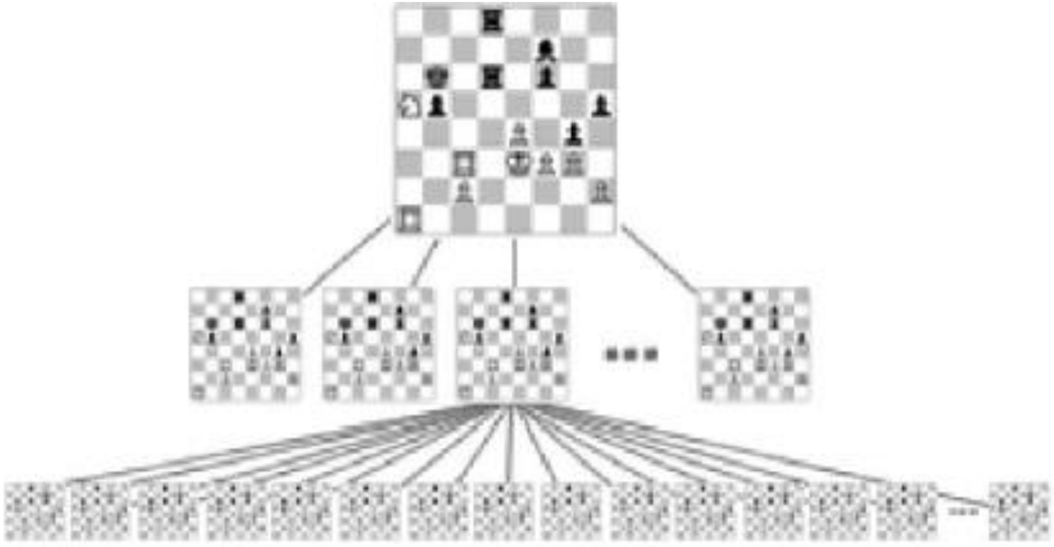
SOME MILESTONES

Is playing chess a matter of intelligence?

vives University of Applied Sciences

Search trees

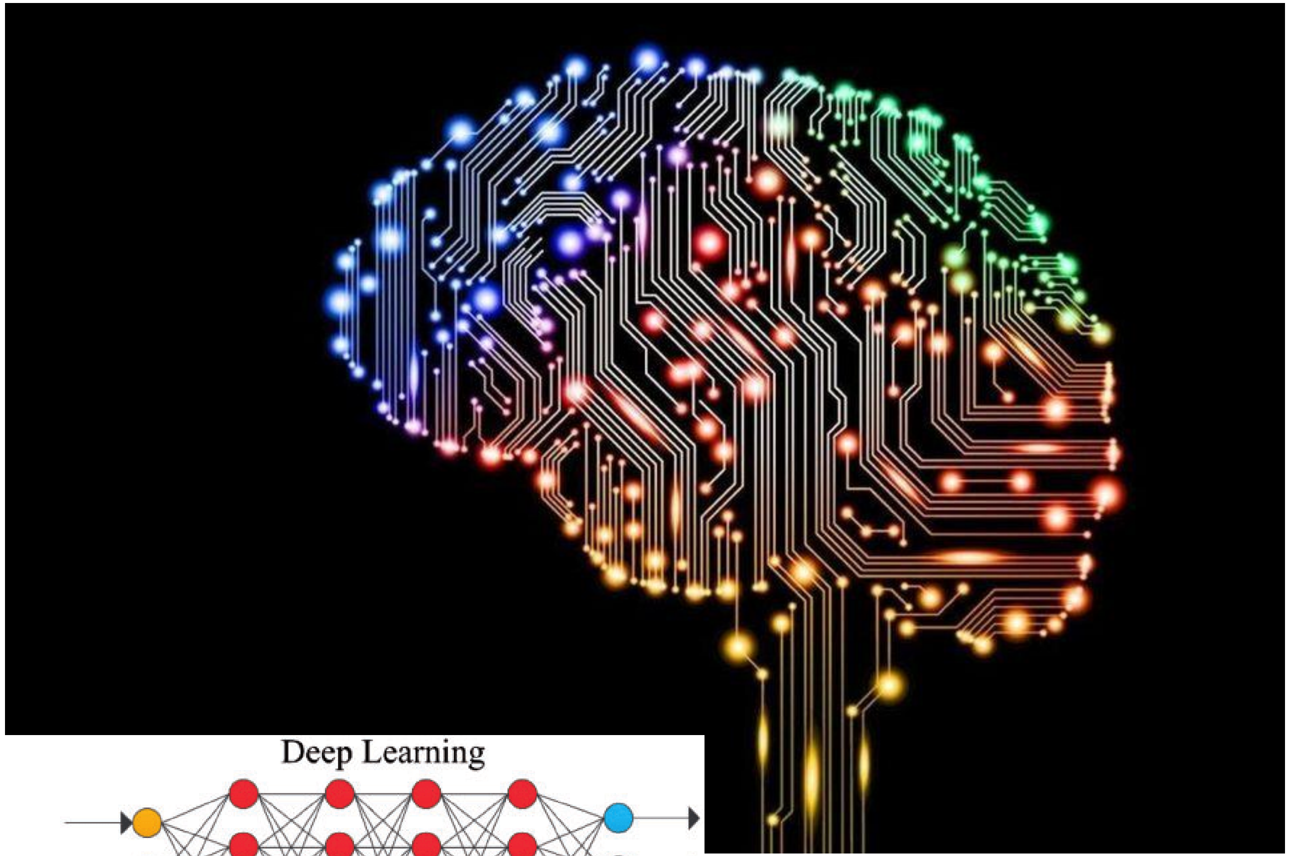


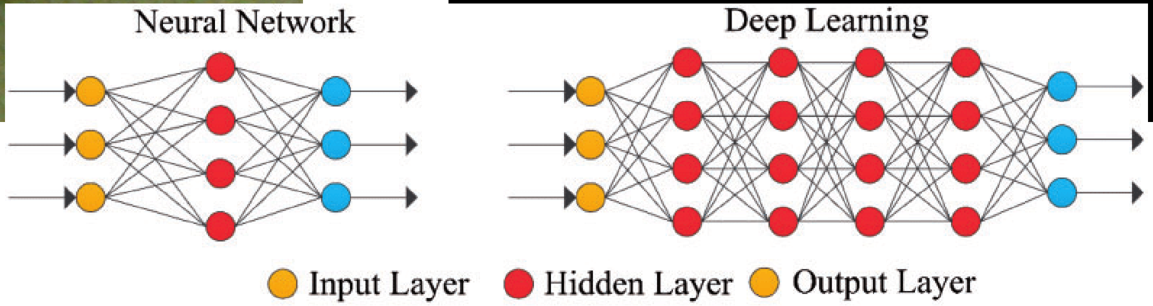


10 years later: Watson

Requires knowledge!

Deep learning & GO





Google Deep Mind: 2016

ChatGPT – 2022 & 2023

- Deep learning based

- New model architecture: transformers

- Language model – does NOT know anything!

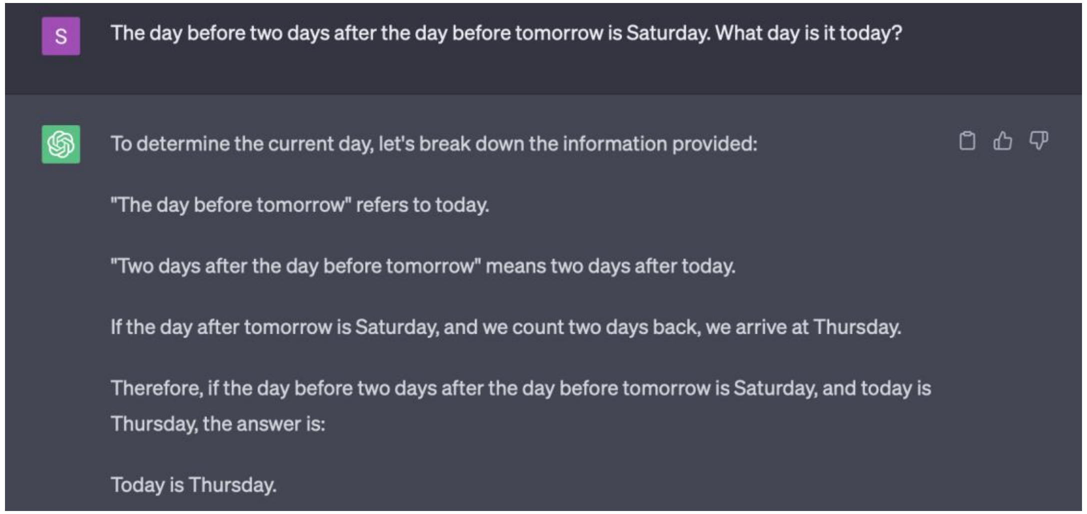
- Predicts the probability that the next word is X given some previous words

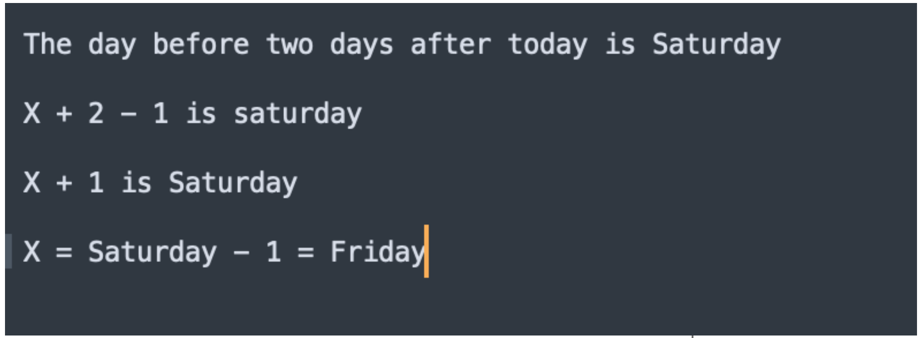
ChatGPT – 2022 & 2023

- The model has several flaws:
  - Bias
  - Racism
  - Wrong answers
  - Solved by adding rules… → degrades to an expert system, a lot of manual programming required!

- The model DOES NOT learn immediately!

- Has NO logic!





Artificial Intelligence: introduction and examples

ABOUT INTELLIGENCE

About intelligence…

- When should we consider a program intelligent?
- When do we consider a (creative) activity of humans to require intelligence?
- Default answers: never? always?
- Most used definition of artificial intelligence:
  - If a task performed by humans requires intelligence, and the software/device is able to perform this task, then the software/device is an artifical intelligence system.

Does numeric computation require intelligence?

For humans?

Xcalc

```
      3921 , 56
  x     73 , 13
-----------------
 286 783 , 68
```

For computers?

➔ Also in the year 1900 ?

When do we consider a program "intelligent"?

To situate the question: two aims

Long term aim:

- develop systems that achieve a level of “intelligence” similar/comparable/better? than that of humans

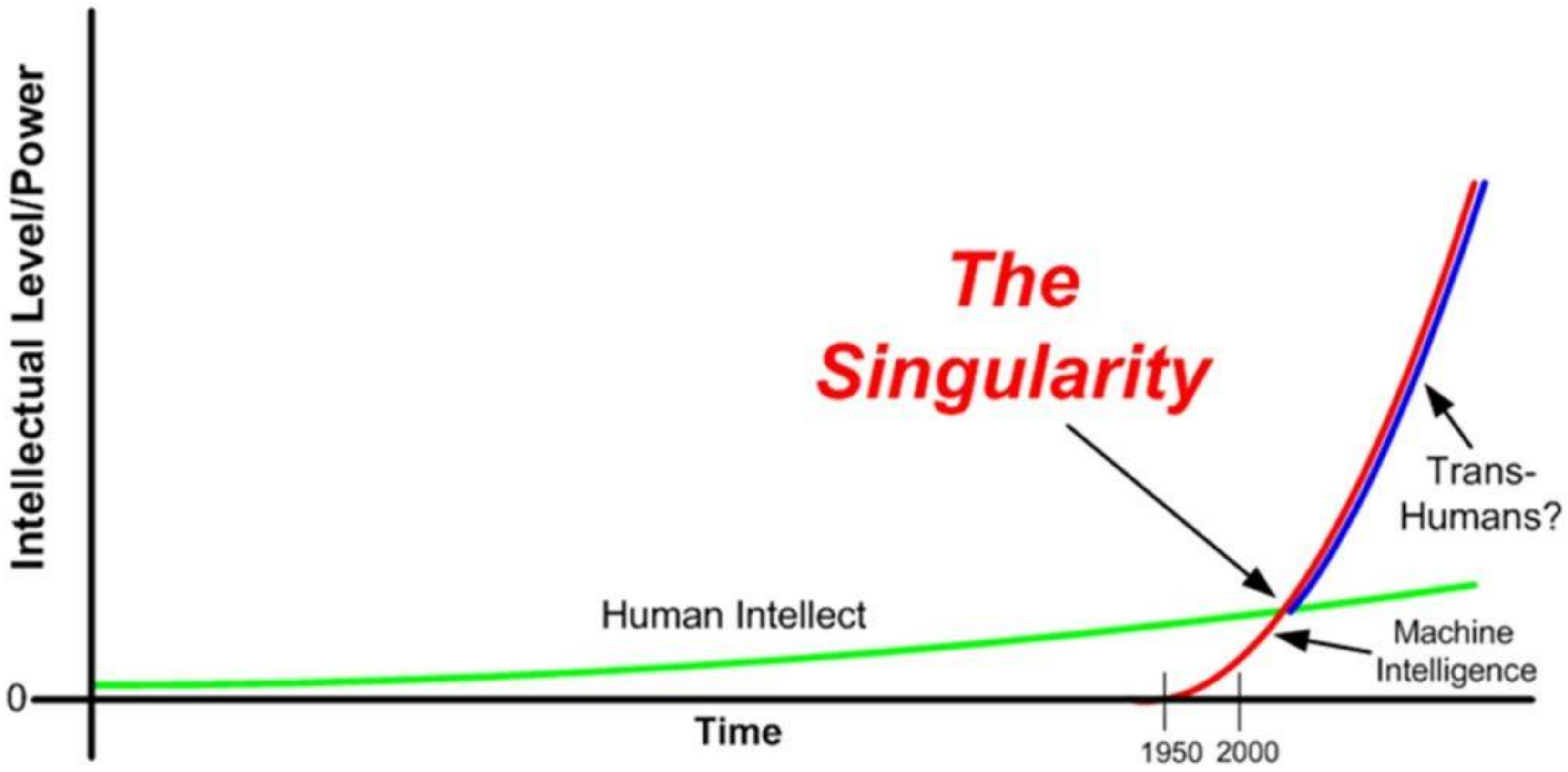
- not achievable in the next 20 to 30 years

Short term aim:

- on specific tasks that seem to require intelligence
- develop systems that achieve a “level of intelligence” similar/comparable/better? than that of humans

- achieved for very many tasks already: deep blue, data mining, computer vision, …

Long term: point of singularity



Intermezzo

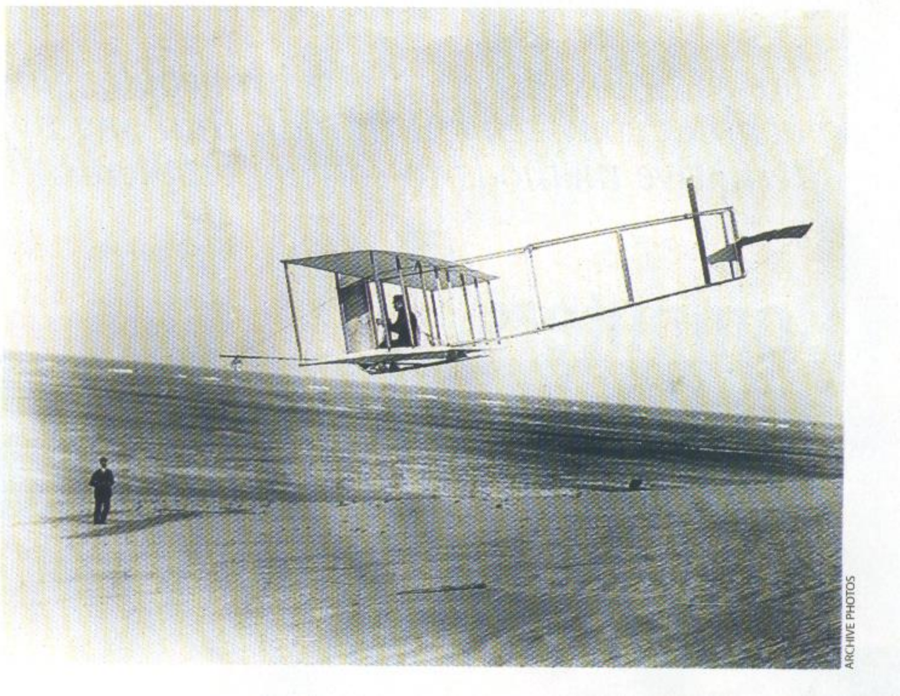
A stupid question:

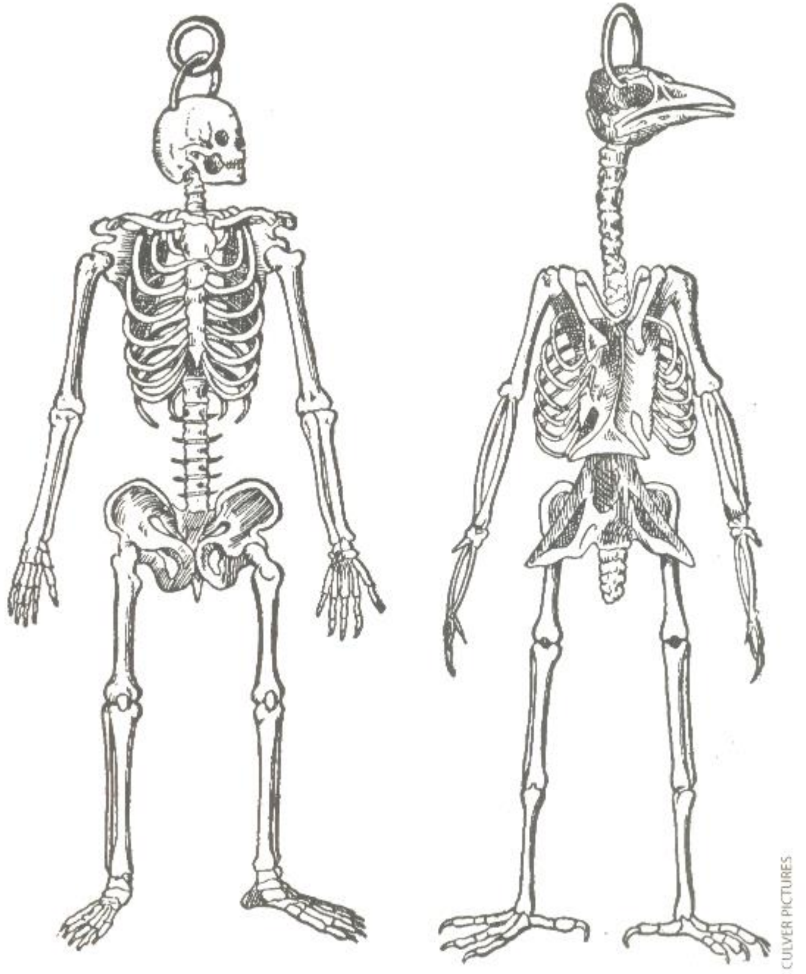
Can people fly?

Short term: reproduction versus simulation

We are not to SIMULATING human intelligence

We are REPRODUCING the effect of intelligence

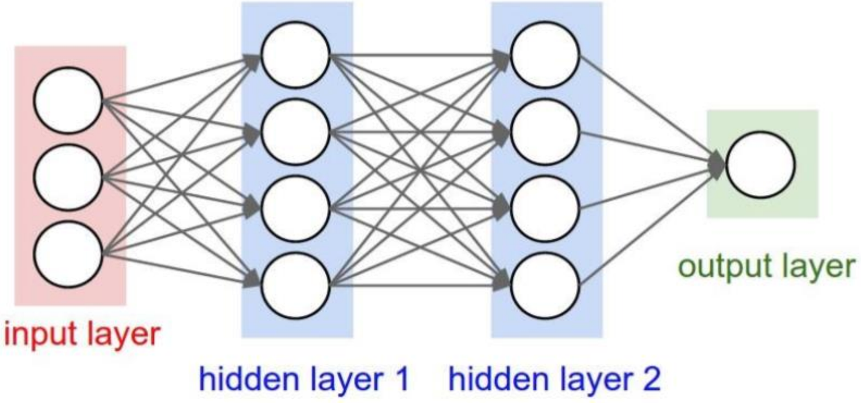




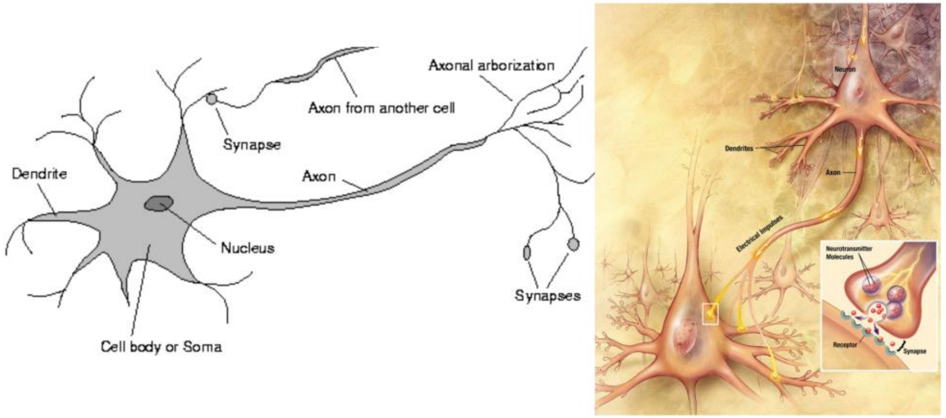
COMPARISON OF SKELETONS of a human and a bird—here taken from a 16th-century manuscript by French naturalist Pierre Belon—examined similarites in anatomy in an attempt to understand how birds can fly.

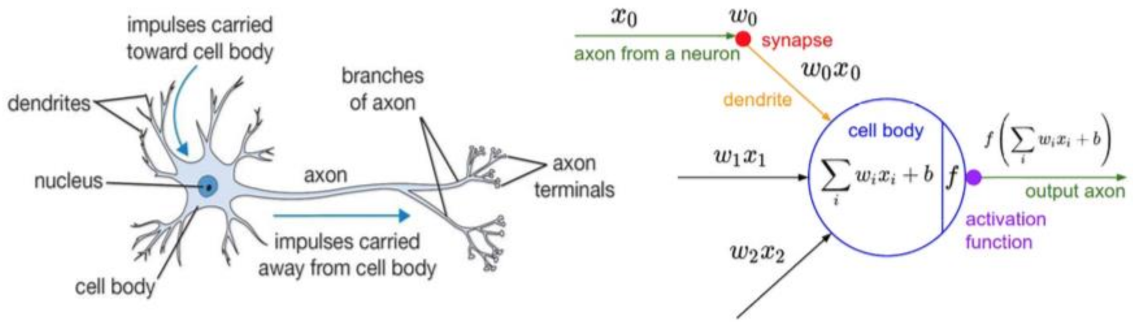
Will AI overrule us?

- Is more or less the question: will deep learning overrule us?



Biology of a Neuron





University
of Applied
Sciences

WHAT IS A.I.?

There are as many definitions as there are practitioners

How would you define it? What is important for a system to be intelligent?

Artificial Intelligence: introduction and examples

A BRIEF HISTORY OF A.I.

vives University of Applied Sciences

Small history & definitions

First used by John McCarthy, 1955

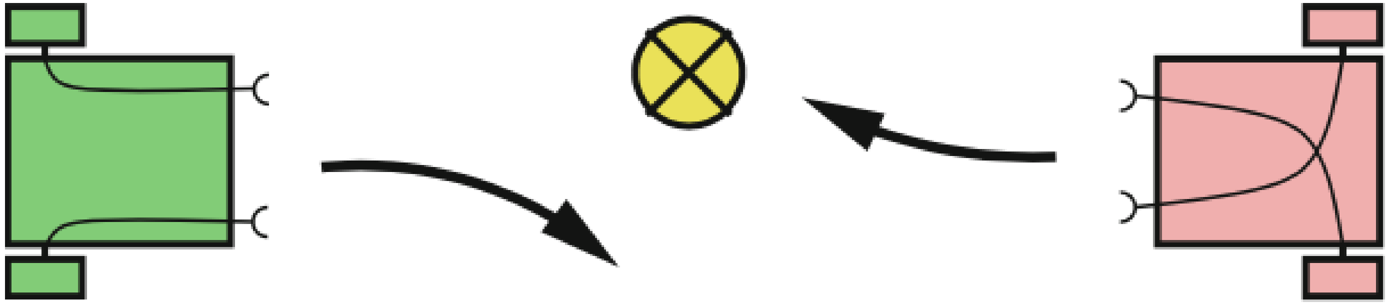
The goal of AI is to develop machines that behave as though they were intelligent

Case:

- 15 robotic vehicles moving around
  - Some vehicles: small groups with little movements
  - Some others: move around peacefully, avoiding collisions
  - Some others: follow a leader
  - Some other: aggressive behaviour
- Question: is this intelligent behaviour?

Small history & definitions

Braitenberg vehicles



So, definition of McCarthy is insufficient

Small history & definitions

Second attempt, definition in Encyclopedia Britannica:

AI is the ability of digital computers or computer controlled robots to solve problems that are normally associated with the higher intellectual processing capabilities of humans …

Is this definition sufficient?

- Computers with large memory to store and retrieve large texts?
- Multiplication

Small history & definitions

Third attempt, Elaine Rich:

Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better.

Long lasting definition?

- Computations?
- Chess?
- Entering a (unknown) room, making an inventory of items, planning and taking actions
  - Field of autonomous robots

History of A.I.

Table 1.1 Milestones in the development of AI from Gödel to today

1931	The Austrian Kurt Gödel shows that in first-order *predicate logic* all true statements are derivable [Göd31a]. In higher-order logics, on the other hand, there are true statements that are unprovable [Göd31b]. (In [Göd31b] Gödel showed that predicate logic extended with the axioms of arithmetic is incomplete.)
1937	Alan Turing points out the limits of intelligent machines with the halting problem [Tur37].
1943	McCulloch and Pitts model *neural networks* and make the connection to propositional logic.
1950	Alan Turing defines machine intelligence with the *Turing test* and writes about learning machines and genetic algorithms [Tur50].
1951	Marvin Minsky develops a neural network machine. With 3000 vacuum tubes he simulates 40 neurons.
1955	Arthur Samuel (IBM) builds a learning checkers program that plays better than its developer [Sam59].
1956	McCarthy organizes a conference in Dartmouth College. Here the name *Artificial Intelligence* was first introduced.
	Newell and Simon of Carnegie Mellon University (CMU) present the *Logic Theorist*, the first symbol-processing computer program [NSS83].
1958	McCarthy invents at MIT (Massachusetts Institute of Technology) the high-level language *LISP*. He writes programs that are capable of modifying themselves.
1959	Gelernter (IBM) builds the Geometry Theorem Prover.
1961	The General Problem Solver (GPS) by Newell and Simon imitates human thought [NS61].
1963	McCarthy founds the AI Lab at Stanford University.
1965	Robinson invents the *resolution calculus* for predicate logic [Rob65] (Sect. 3.5).
1966	Weizenbaum's program Eliza carries out dialog with people in natural language [Wei66] (Sect. 1.1.2).
1969	Minsky and Papert show in their book *Perceptrons* that the perceptron, a very simple neural network, can only represent linear functions [MP69] (Sect. 1.1.2).
1972	French scientist Alain Colmerauer invents the logic programming language *PROLOG* (Chap. 5).
	British physician de Dombal develops an *expert system* for diagnosis of acute abdominal pain [dDLS+72]. It goes unnoticed in the mainstream AI community of the time (Sect. 7.3).
1976	Shortliffe and Buchanan develop MYCIN, an expert system for diagnosis of infectious diseases, which is capable of dealing with uncertainty (Chap. 7).
1981	Japan begins, at great expense, the "Fifth Generation Project" with the goal of building a powerful PROLOG machine.
1982	R1, the expert system for configuring computers, saves Digital Equipment Corporation 40 million dollars per year [McD82].
1986	Renaissance of neural networks through, among others, Rumelhart, Hinton and Sejnowski [RM86]. The system Nettalk learns to read texts aloud [SR86] (Chap. 9).
1990	Pearl [Pea88], Cheeseman [Che85], Whittaker, Spiegelhalter bring probability theory into AI with *Bayesian networks* (Sect. 7.4). Multi-agent systems become popular.

(continued)

History of A.I.

Table 1.1 (continued)

1992	Tesauros TD-gammon program demonstrates the advantages of reinforcement learning.
1993	Worldwide *RoboCup* initiative to build soccer-playing autonomous robots [Roba].
1995	From statistical learning theory, Vapnik develops support vector machines, which are very important today.
1997	IBM's chess computer Deep Blue defeats the chess world champion Gary Kasparov.
	First international RoboCup competition in Japan.
2003	The robots in RoboCup demonstrate impressively what AI and robotics are capable of achieving.
2006	Service robotics becomes a major AI research area.
2009	First Google self-driving car drives on the California freeway.
2010	Autonomous robots begin to improve their behavior through learning.
2011	IBM's "Watson" beats two human champions on the television game show "Jeopardy!". Watson understands natural language and can answer difficult questions very quickly (Sect. 1.4).
2015	Daimler premiers the first autonomous truck on the Autobahn.
	Google self-driving cars have driven over one million miles and operate within cities.
	Deep learning (Sect. 11.9) enables very good image classification.
	Paintings in the style of the Old Masters can be automatically generated with deep learning. AI becomes creative!
2016	The Go program AlphaGo by Google DeepMind [SHM+16] beats the European champion 5:0 in January and Korean Lee Sedol, one of the world's best Go players, 4:1 in March. Deep learning techniques applied to pattern recognition, as well as reinforcement learning and Monte Carlo tree search lead to this success.

History of A.I.

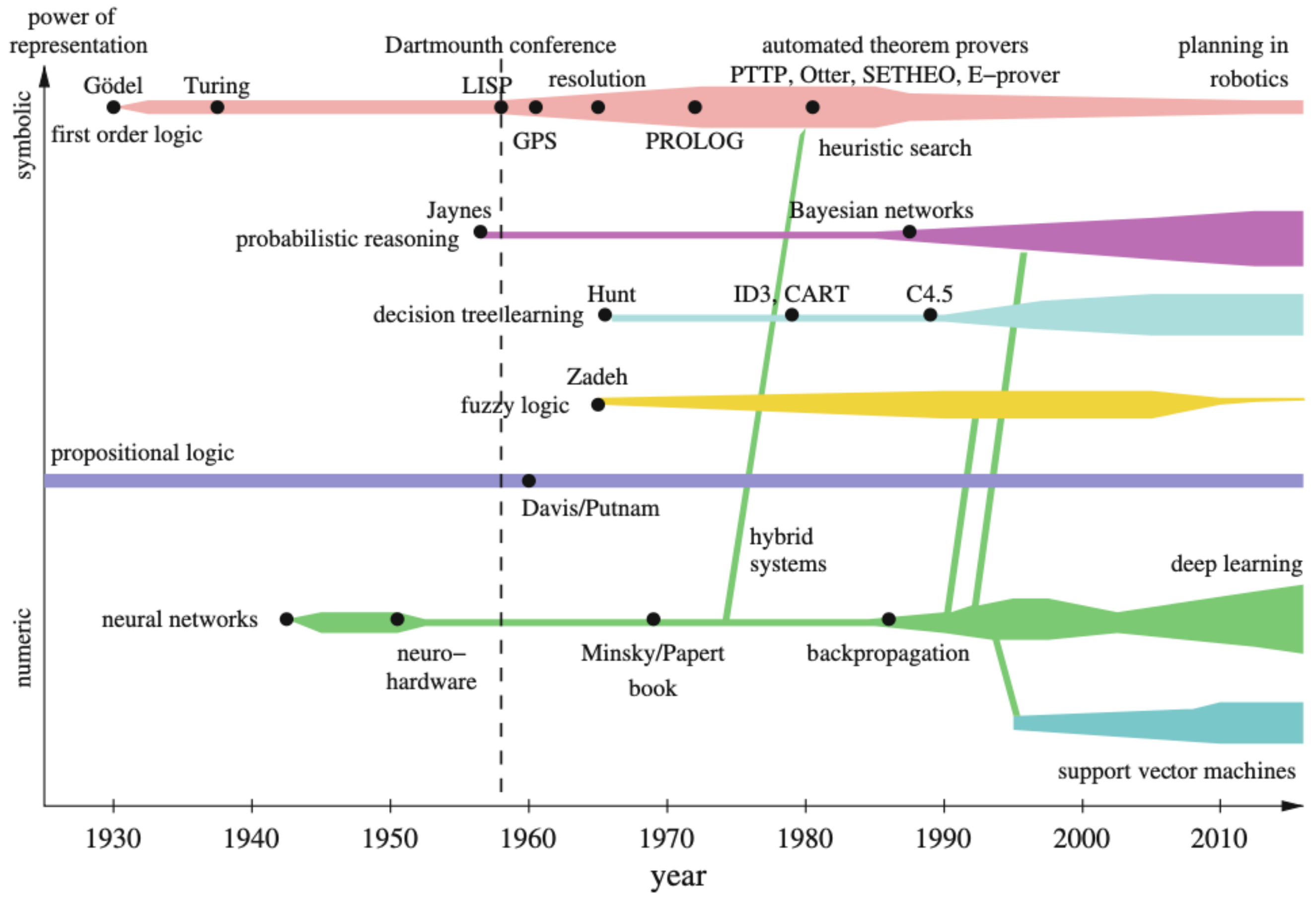


Fig. 1.3 History of the various AI areas. The width of the *bars* indicates prevalence of the method's use

What is A.I.?

Artificial intelligence is technology that appears to **emulate human performance** typically by **learning**, coming to its own conclusions, appearing to **understand** complex content, engaging in natural dialogs with people, **enhancing human cognitive performance** (also known as cognitive computing) or **replacing** people on execution of **nonroutine** (but repetitive!) tasks.

Applications include **autonomous** vehicles, **automatic** speech recognition and **generation** and **detecting** novel concepts and abstractions (useful for detecting potential new risks and aiding humans quickly understand very large bodies of everchanging information).

Intelligence in A.I.

- Ability to interact with the world (speech, vision, motion, manipulation)

- Ability to model the world and to reason about it

- And in the best case: ability to learn and to adapt

AI – ML - DL

- Artificial Intelligence (AI):

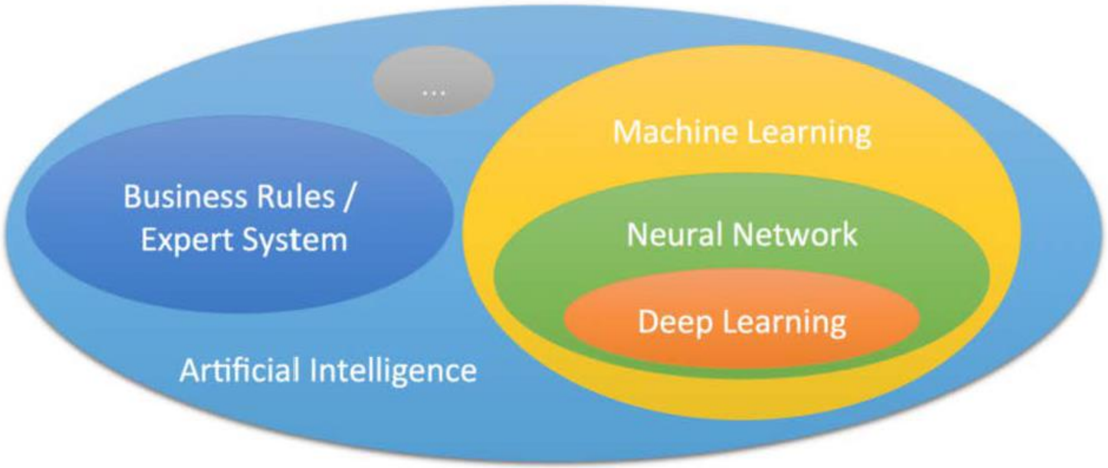
  "The set of all tasks in which a computer can make decisions."

- Machine Learning (ML):

  "The set of all tasks in which a computer can make decisions *based on data*."

- Deep Learning (DL):

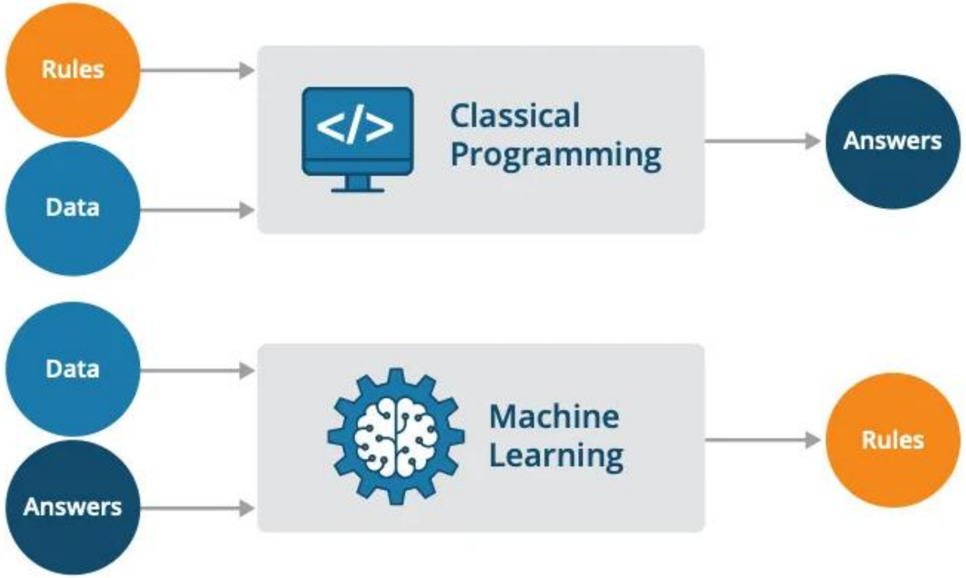
  "The field of machine learning that uses certain objects called *neural networks*."

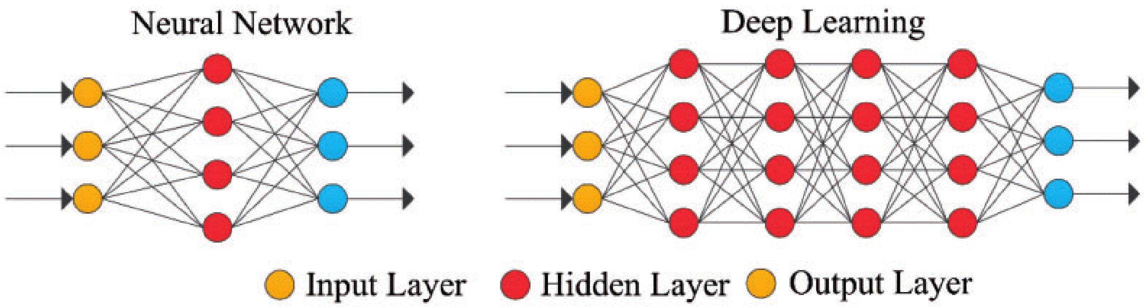


AI - ML

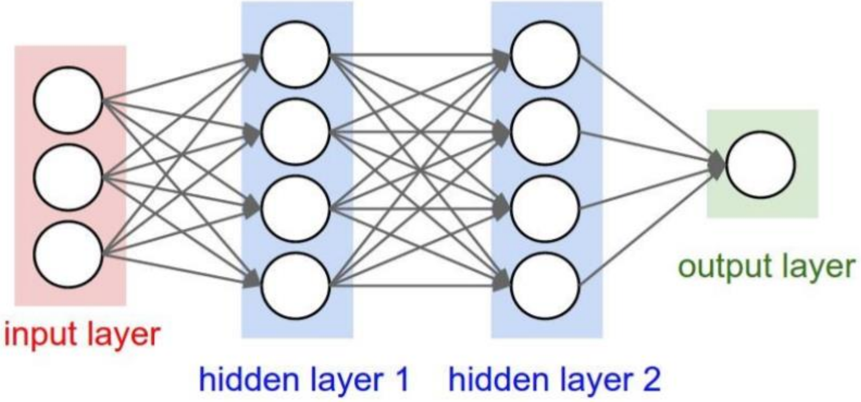
Issue with ML:

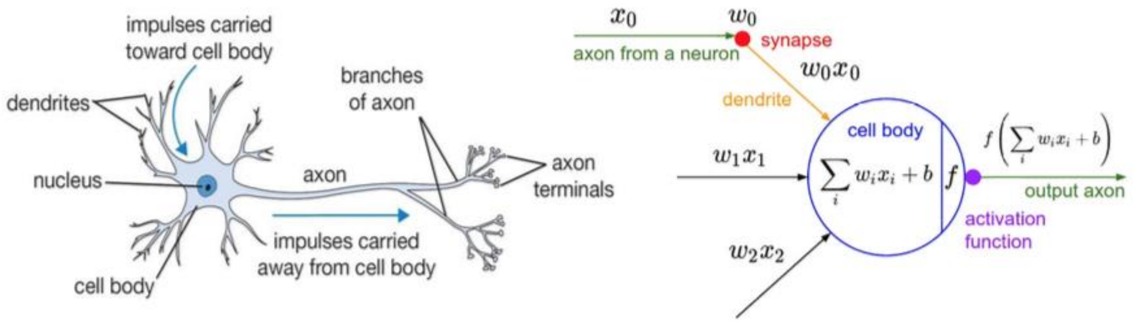
- Not always interpretable
- Needs (a lot of) data





AI - DL





vives University of Applied Sciences

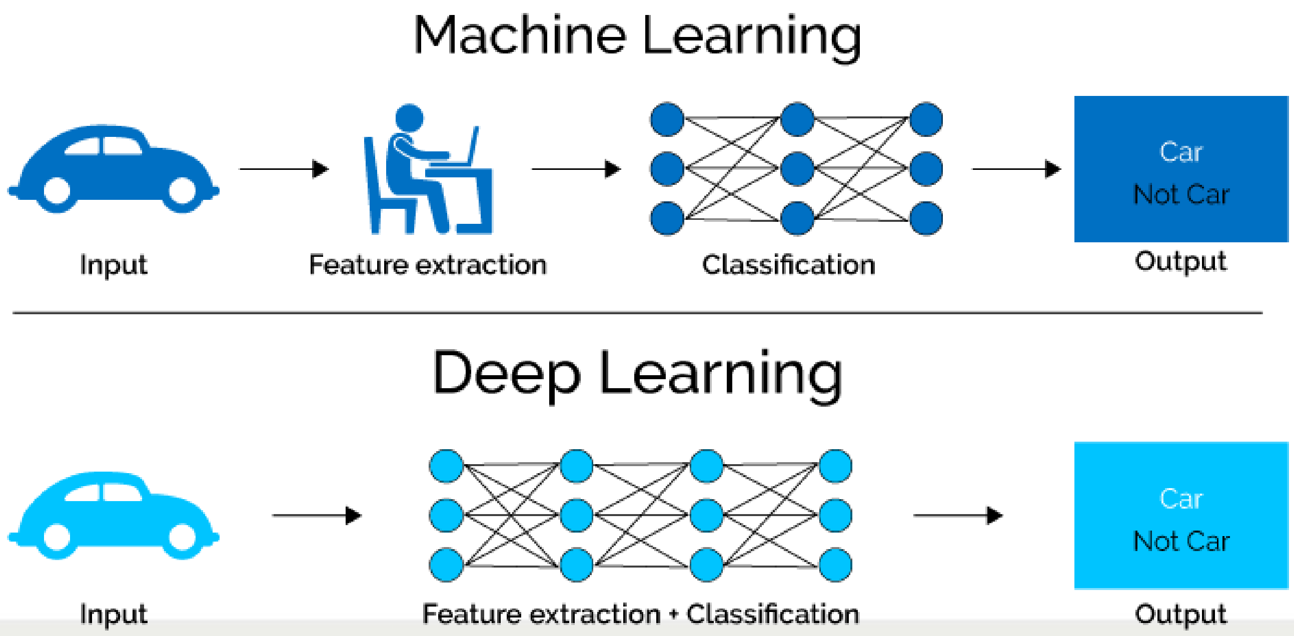
What is deep learning? – foto brein!

Subfield of ML for learning representations of data.

Exceptional effective at learning patterns.

Utilizes learning algorithms that derive meaning out of data by using a hierarchy of multiple layers that mimic the neural networks of our brain.

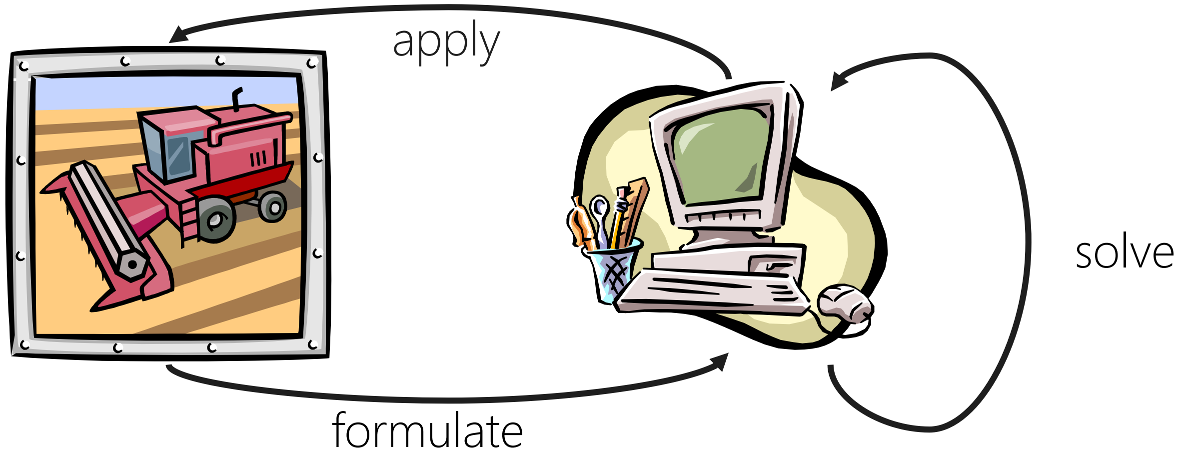
If you provide the system tons of information, it begins to understand it and respond in useful ways.



https://www.xenonstack.com/blog/static/public/uploads/media/machine-learning-vs-deep-learning.png

What is AI? For me...

Computers/algorithms making decisions/predictions in real-world problems



A real world example:
the nurse rostering problem (NRP)

	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S	S
A	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH	DH		DH	DH	DH	DH	DH	DH	DH		
B	L	L	V	V	V					D	D		L	L	L	D	V	V				[illegible]	[illegible]		D	[illegible]	[illegible]	[illegible]
C	D	D			L	L	L		[illegible]	[illegible]					V	V	D		[illegible]	[illegible]	[illegible]		V	D	L	L	V	V
D	[illegible]	[illegible]	[illegible]	[illegible]		D	D	V	V	V		D	V	V				L	L				L	L		D	D	D
E			L	L	D	V	V	D	D		L	L			[illegible]	[illegible]		D	DH				D	V	V	V	DH	DH
F	V	V	D	D				L	L	L	[illegible]	[illegible]	[illegible]	[illegible]		L	L		V	V	V	D		[illegible]	[illegible]			
G					[illegible]	[illegible]	[illegible]	[illegible]											D	D	D	V					L	L
H											V	V	D	D	D		[illegible]	[illegible]		L	L	L						

Problem:

assignments of shifts to nurses with a certain qualification considering a number of constraints

AI – example application 1 - nurse rostering

- **Small example:**
  - 4 shifts to schedule
  - 10 possible nurses
  - 7 days

- **Number of variations on 1 day = 5040**
  - For 1 week: number of possible rosters = 5040^7 = 8,260641125390352e25
  - For 1 month: number of possible rosters = 5040^{30} = 1,182813011613388e111

AI – example application 2 - transport for disabled people

- **Goal:**
  - Drawing up timetables for collective transport for people with disabilities from home to reception centre

- **Problem:**
  - Clients
    - Type of wheelchair
    - Earliest and latest pickup and delivery time
    - Pickup location
  - Depots = day care centres
    - Earliest and latest arrival and departure time
  - Vehicles
    - Capacities
    - Different configurations

vives University of Applied Sciences

AI – example application 2 - transport for disabled people

Credits: Tellez et al. 2018

x_{ij}^{kc} is a binary variable which is equal to 1 if vehicle $k \in \mathscr{K}$ uses arc $(i,j) \in \mathscr{A}$ with configuration $c \in \mathscr{C}_k$, and 0 otherwise,
z_i^k is a binary variable which is equal to 1 if vehicle $k \in \mathscr{K}$ is reconfigured at node $i \in \mathscr{P} \cup \mathscr{D}$, and 0 otherwise.
$l_{i,u}^k$ is an integer variable representing the number of users of type $u \in \mathscr{U}$ in vehicle $k \in \mathscr{K}$ after visiting node $i \in \mathscr{V}$,
w_i^k is a continuous variable representing the time of service of vehicle $k \in \mathscr{K}$ at node $i \in \mathscr{V}$.

The FSM-DARP-RC can then be formulated with the following mixed integer program:

$$\min \sum_{k\in\mathscr{K}} f^k \sum_{i\in\mathscr{P}} \sum_{c\in\mathscr{C}^k} x_{o_k^+ i}^{kc} + \alpha \sum_{k\in\mathscr{K}} (w_{o_k^-}^k - w_{o_k^+}^k) + \sum_{k\in\mathscr{K}} \sum_{c\in\mathscr{C}^k} \sum_{(i,j)\in\mathscr{A}} \gamma^k \Delta_{ij} x_{ij}^{kc} \tag{1}$$

s.t.

$$\sum_{c\in\mathscr{C}^k} \sum_{(p_r,j)\in\mathscr{A}} x_{p_r j}^{kc} - \sum_{c\in\mathscr{C}^k} \sum_{(j,d_r)\in\mathscr{A}} x_{j d_r}^{kc} = 0 \quad \forall\, r \in \mathscr{R},\ k \in \mathscr{K} \tag{2}$$

$$\sum_{c\in\mathscr{C}^k} \sum_{k\in\mathscr{K}} \sum_{(j,p_r)\in\mathscr{A}} x_{j p_r}^{kc} = 1 \quad \forall\, r \in \mathscr{R} \tag{3}$$

$$\sum_{c\in\mathscr{C}^k} \sum_{(j,i)\in\mathscr{A}} x_{ji}^{kc} - \sum_{c\in\mathscr{C}^k} \sum_{(i,j)\in\mathscr{A}} x_{ij}^{kc} = 0 \quad \forall\, i \in \mathscr{P} \cup \mathscr{D},\ k \in \mathscr{K} \tag{4}$$

$$\sum_{c\in\mathscr{C}^k} \sum_{i\in\mathscr{P}} x_{o_k^+ i}^{kc} - \sum_{c\in\mathscr{C}^k} \sum_{i\in\mathscr{D}} x_{i o_k^-}^{kc} = 0 \quad \forall\, k \in \mathscr{K} \tag{5}$$

$$w_j^k \geqslant w_i^k + t_{ij} + s_i - M_{ij}\Big(1 - \sum_{c\in\mathscr{C}^k} x_{ij}^{kc}\Big) \quad \forall\, (i,j) \in \mathscr{A},\ k \in \mathscr{K} \tag{6}$$

$$a_i \leqslant w_i^k \leqslant b_i \quad \forall\, i \in \mathscr{V},\ k \in \mathscr{K} \tag{7}$$

$$w_{p_r}^k + s_{p_r} + t_{p_r d_r} \leqslant w_{d_r}^k \quad \forall\, r \in \mathscr{R},\ k \in \mathscr{K} \tag{8}$$

$$l_{j,u}^k \geqslant l_{i,u}^k + \phi_{ju} - \overline{Q}_u^k \left(1 - \sum_{c\in\mathscr{C}^k} x_{ij}^{kc}\right) \quad \forall\, (i,j) \in \mathscr{A},\ u \in \mathscr{U},\ k \in \mathscr{K} \tag{9}$$

$$l_{i,u}^k \leqslant \sum_{c\in\mathscr{C}^k} \sum_{(i,j)\in\mathscr{A}} Q_u^{kc} x_{ij}^{kc} \quad \forall\, i \in \mathscr{P},\ u \in \mathscr{U},\ k \in \mathscr{K} \tag{10}$$

$$\sum_{c'\in\mathscr{C}^k/\{c\}} \sum_{(j,i)\in\mathscr{A}} x_{ji}^{kc'} + \sum_{(i,j)\in\mathscr{A}} x_{ij}^{kc} \leqslant 1 + z_i^k \quad \forall\, i \in \mathscr{P} \cup \mathscr{D},\ k \in \mathscr{K},\ c \in \mathscr{C}^k \tag{11}$$

$$\sum_{i\in\mathscr{P}\cup\mathscr{D}} z_i^k \leqslant \overline{R} \quad \forall\, k \in \mathscr{K} \tag{12}$$

$$w_{d_r}^k - w_{p_r}^k - s_{p_r} \leqslant \overline{T}_r \quad \forall\, r \in \mathscr{R},\ k \in \mathscr{K} \tag{13}$$

$$w_{o_k^-}^k - w_{o_k^+}^k \leqslant \overline{T} \quad \forall\, k \in \mathscr{K} \tag{14}$$

$$x_{ij}^{kc} \in \{0,1\} \quad \forall\, (i,j) \in \mathscr{A},\ k \in \mathscr{K} \tag{15}$$

$$l_{i,u}^k \in \mathbb{Z}^+ \quad \forall\, i \in \mathscr{V},\ u \in \mathscr{U},\ k \in \mathscr{K} \tag{16}$$

$$w_i^k \in \mathbb{R}^+ \quad \forall\, i \in \mathscr{V},\ k \in \mathscr{K} \tag{17}$$

$$z_i^k \in \{0,1\} \quad \forall\, i \in \mathscr{P} \cup \mathscr{D},\ k \in \mathscr{K}. \tag{18}$$

AI – example application 2 - transport for disabled people

<u>How to solve?</u>

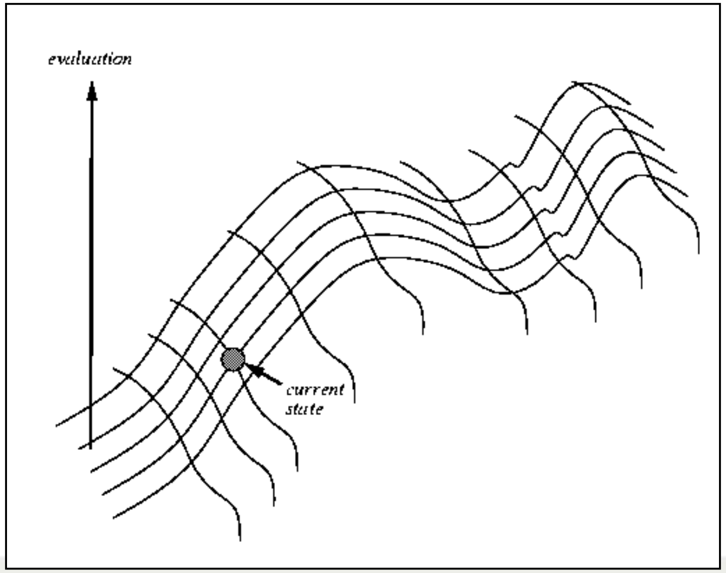
- Integer programming:
  - Modelling is important!
  - Excellent commercial solvers:
    - CPLEX, GUROBI, …
    - can be expensive
  - Great quality open-source solvers:
    - COIN-OR, …
    - Not always stable
  - In general:
    - Can take a lot of time
    - Calculation power

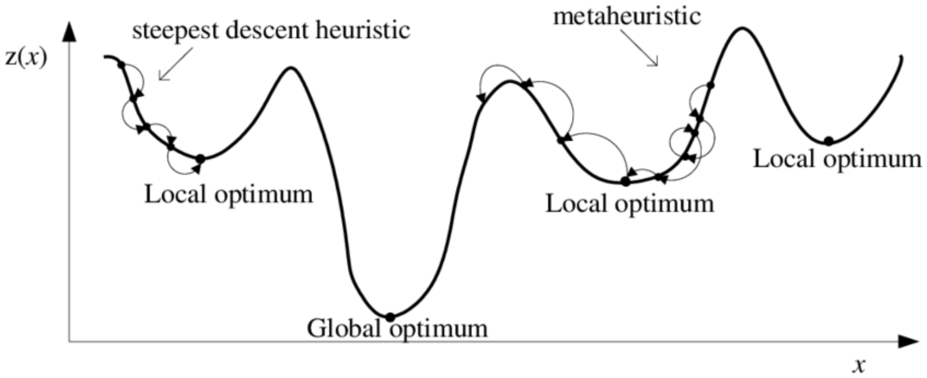
AI – example application 2 - transport for disabled people

How to solve?

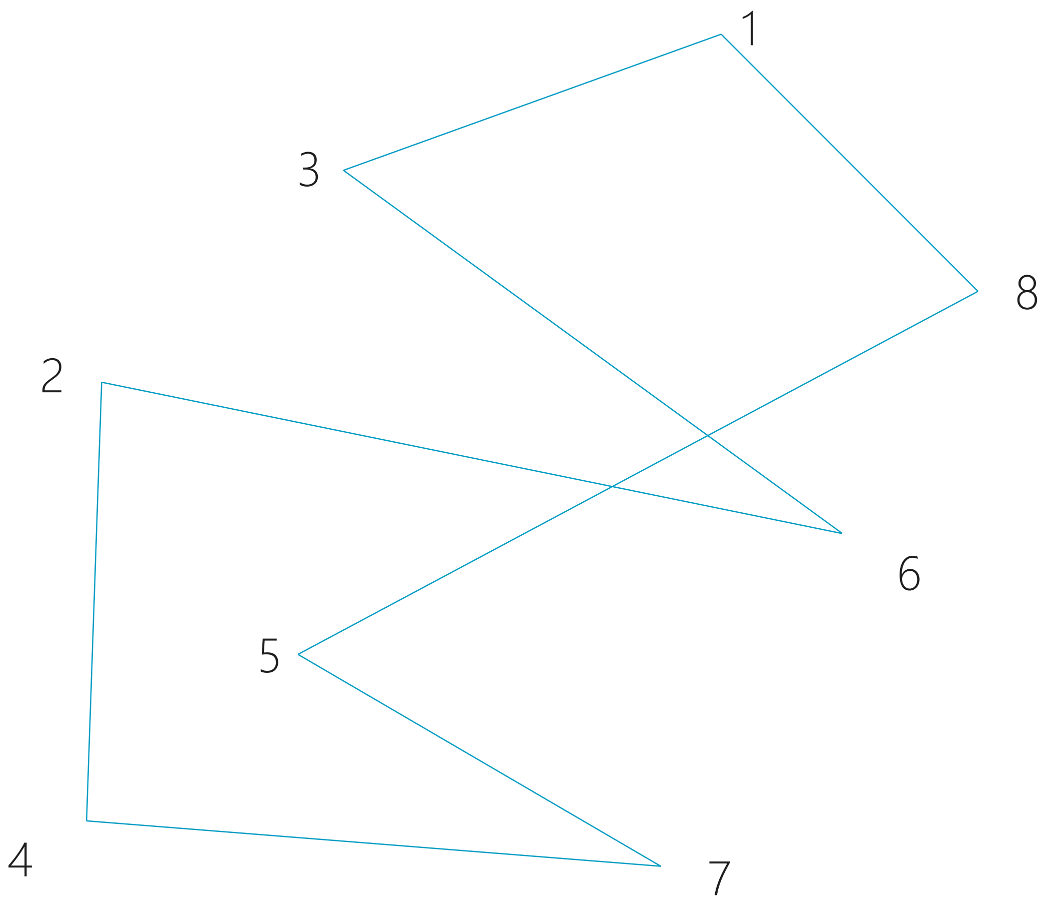
- Use metaheuristics:
  - Generate a start solution
  - "Improve" the solution, by making "specific changes" to the solution
  - Stop if no better solution can be found

- Often used:
  - Hill climbing
  - Steepest descent
  - Tabu Search
  - Genetic algorithms
  - ALNS

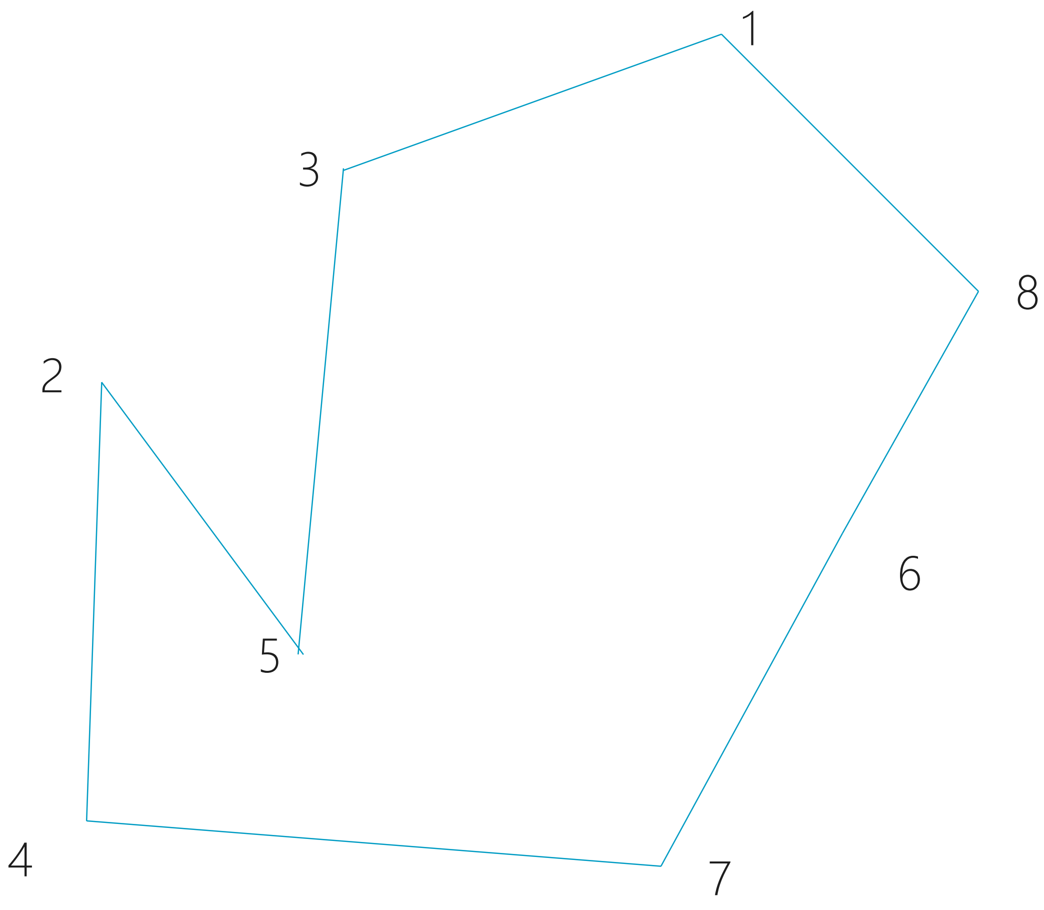




AI – example application 2 - transport for disabled people



AI – example application 2 - transport for disabled people



ALNS – adaptive large neighbourhood search

Solution = ALNS

- Large = "destroy" and "restore" larger parts of the solution

2 phases:

- Remove/destroy-fase
- Repair-fase

Various remove and repair operators possible

Good results: able to timetable 100 requests for 6 buses in less than 1 hour!

AI – Other applications – ML & DL

AI in biomedical applications:

- Early detection of kidny failure after ICU → survival analysis
- Drug discovery based on machine learning

AI in education:

- (Early) detection of dropout based on moodle data
- Recommendation systems for personalised learning

AI in news:

- "Safe" recommendation enginges for news sites

AI in industry:

- Early detection of wear in a machine → survival analysis

vives University of Applied Sciences

AI for optimisation problems

Questions?

Contact details:

Dr. Stefaan Haspeslagh

Liaison Officer AI, KU Leuven – Vives

Doorniksesteenweg 145 – 8500 Kortrijk - Belgium

stefaan.haspeslagh@vives.be

stefaan.haspeslagh@kuleuven.be

vives University of Applied Sciences