INESC TEC

INSTITUTO DE ENGENHARIA
DE SISTEMAS E COMPUTADORES,
TECNOLOGIA E CIÊNCIA

Learning How Machines Learn: An Introduction to Artificial Intelligence and Machine Learning

Tertúlia | Universidade Católica Portuguesa – Católica Porto Business School

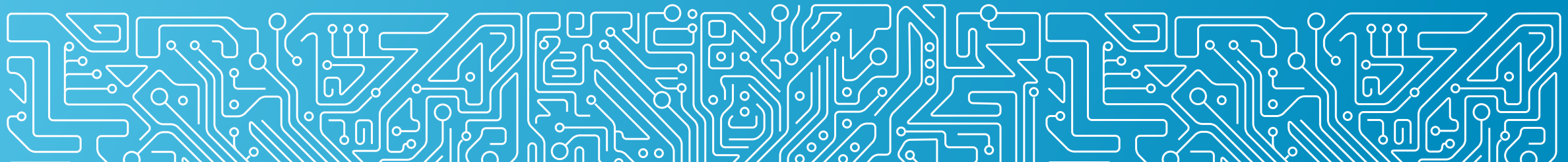
July 7, 2025 | Porto, Portugal

Tiago Filipe Sousa Gonçalves (tiago.f.goncalves@inesctec.pt)

INESC TEC

Preface

History, context and motivation



What is Artificial Intelligence (AI)? What can we tell you about its history?

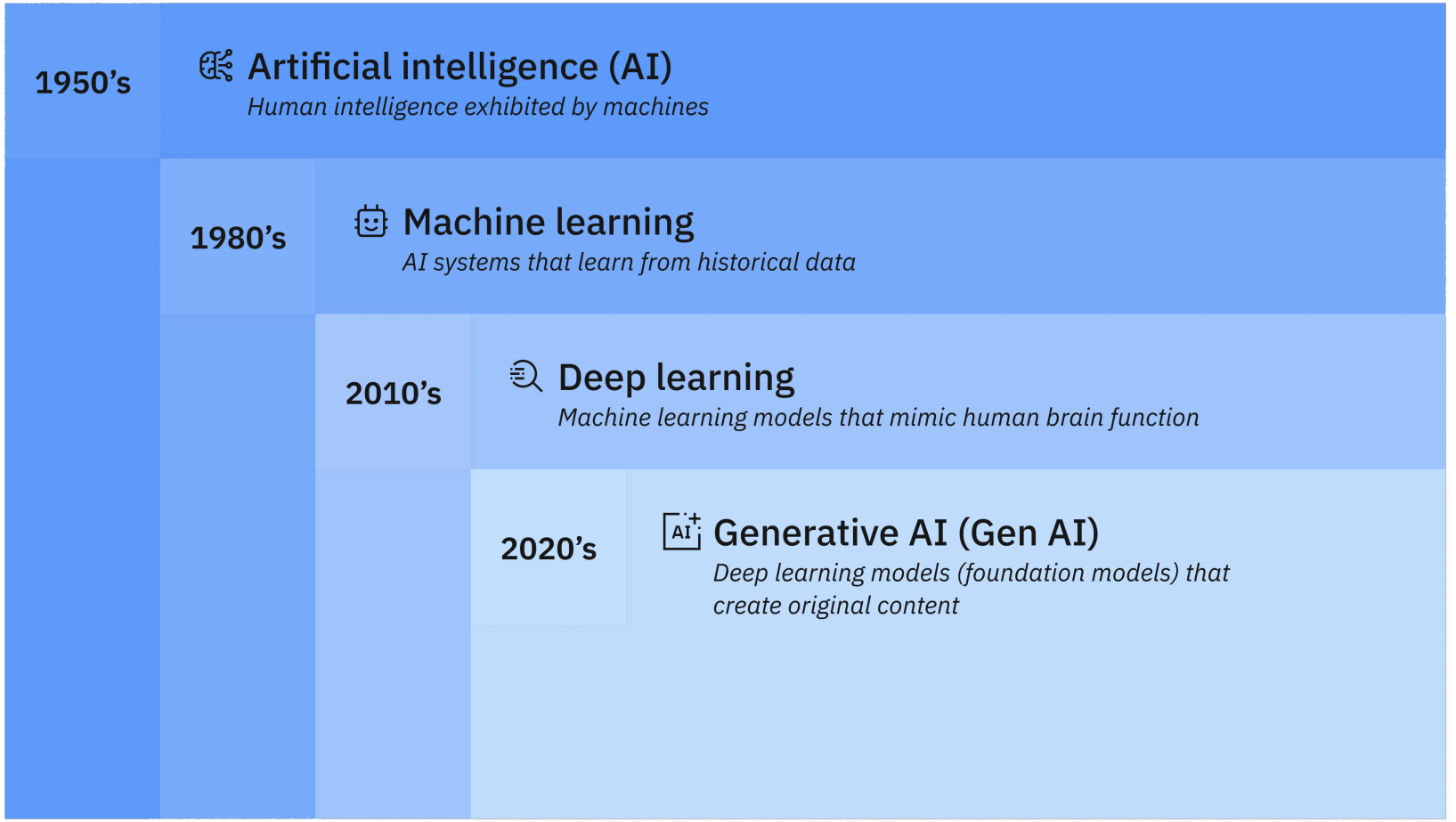
Luger and Stubblefield consider AI as "**the branch of computer science** that is concerned with the **automation** of **intelligent behaviour**"[1]

Let us highlight some interesting dates in the history of AI:

- 1950: Alan Turing published "Computer Machinery and Intelligence" which proposed The Turing Test[2]
- 1955: John McCarthy held a workshop at Dartmouth on "artificial intelligence" which is the first use of the term[3]
- 1997: IBM's Deep Blue beat the Chess World Champion Garry Kasparov[4]
- 2016: Google's DeepMind AlphaGo beat the Go World Champion Lee Sedol[5]
- 2020: Google's DeepMind published a paper suggesting that "its model was able to spot cancer in de-identified screening mammograms with fewer false positives and false negatives than experts"[6]

Sources: [1] Luger and Stubblefield (1993), [2] Turing (1950), [3] Computer History Museum, [4] The Conversation, [5] DeepMind, [6] McKinney al. (2020)

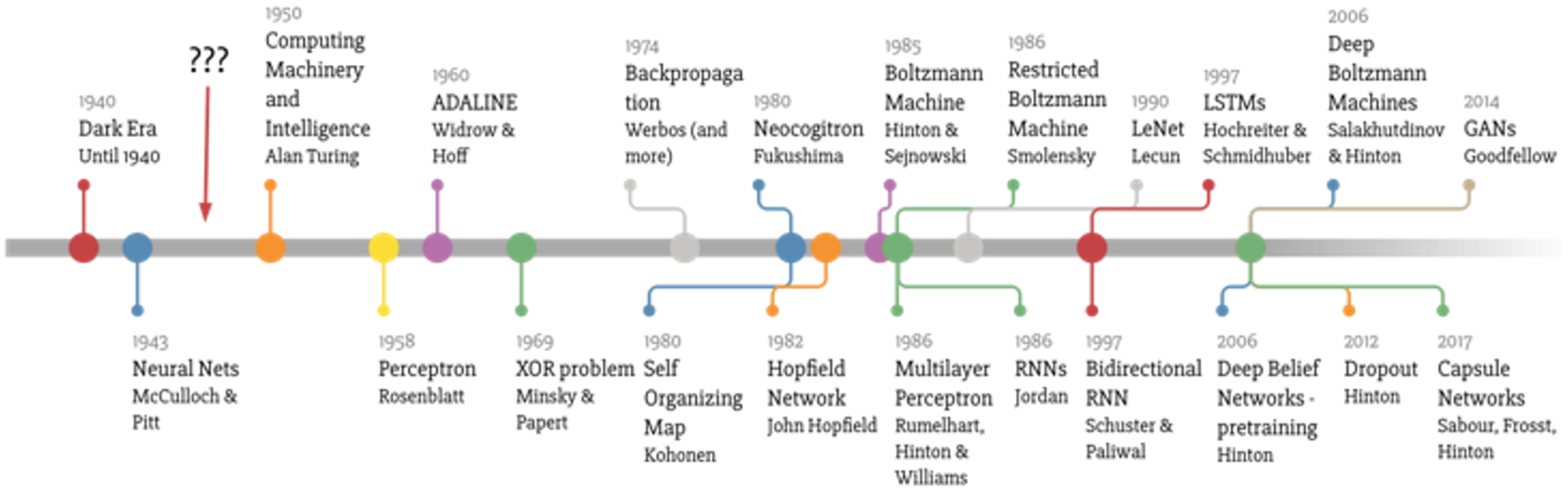
We are entering a new Era of AI



Sources: IBM

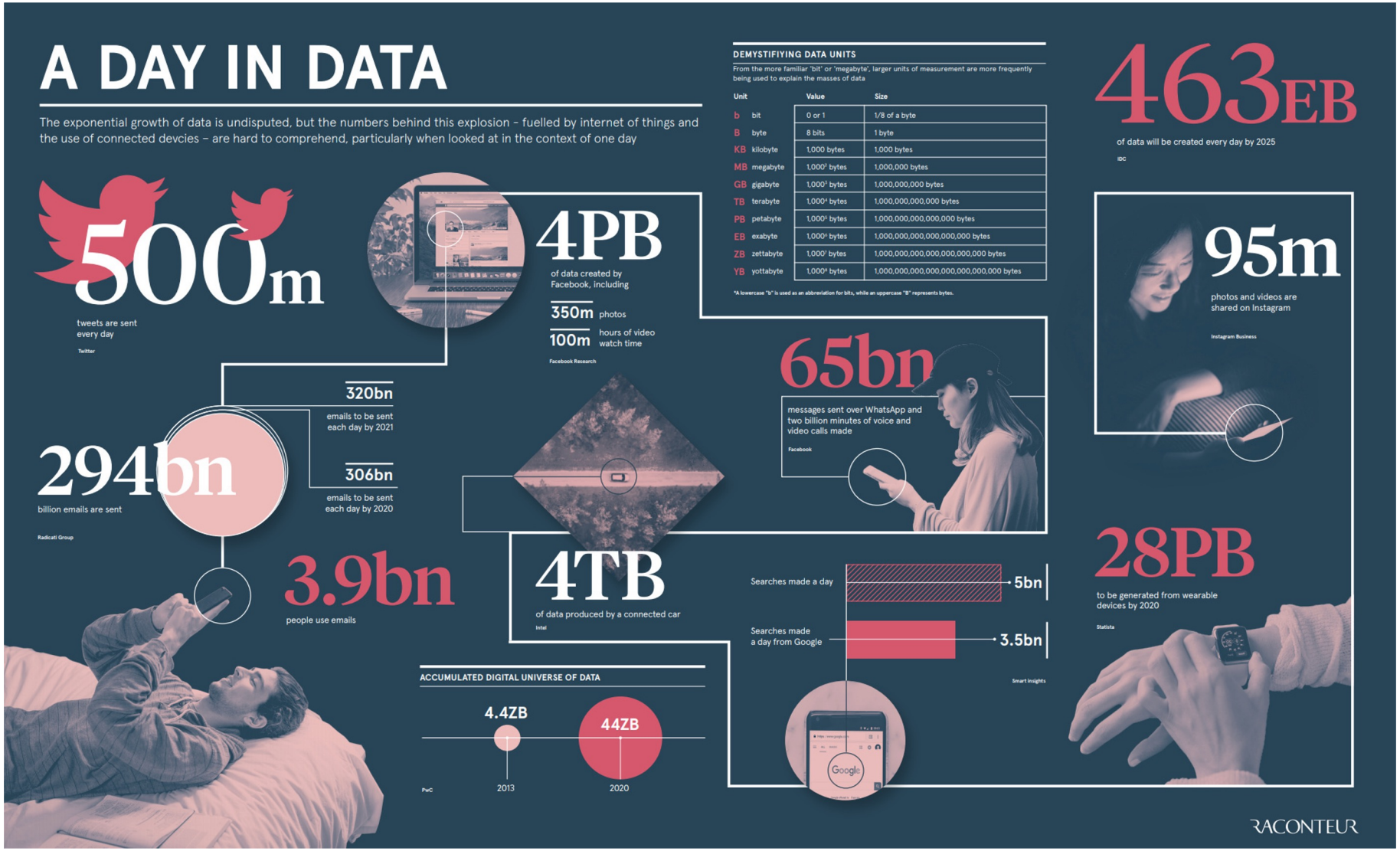
We have more computational power than ever

The development of powerful computer processing units (**CPUs**) and the leveraging of the graphical processing units (**GPUs**) for computation allowed the **training of deep and complex algorithms** in "**human time**"[1,2]



Sources: [1] Towards Data Science, [2] NVIDIA

We generate large quantities of data every day

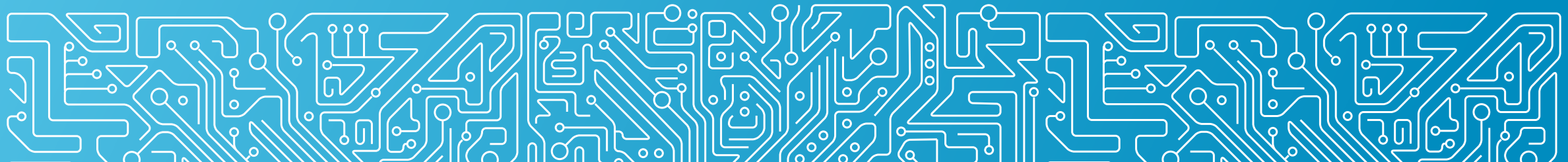


Sources: Raconteur

INESC TEC

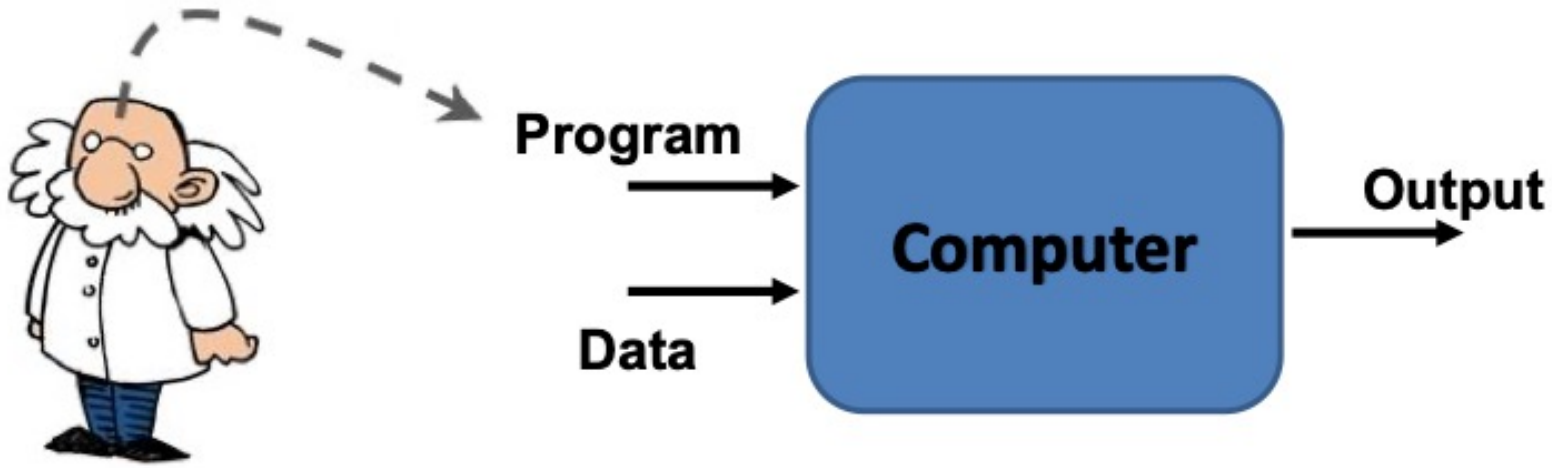
How do machines learn?

Algorithms and how they work



In the beginning, we had rules based on expert knowledge

We used “**if-then-else**” logic to build programs that **simulated** the decision-making process of a **human expert**



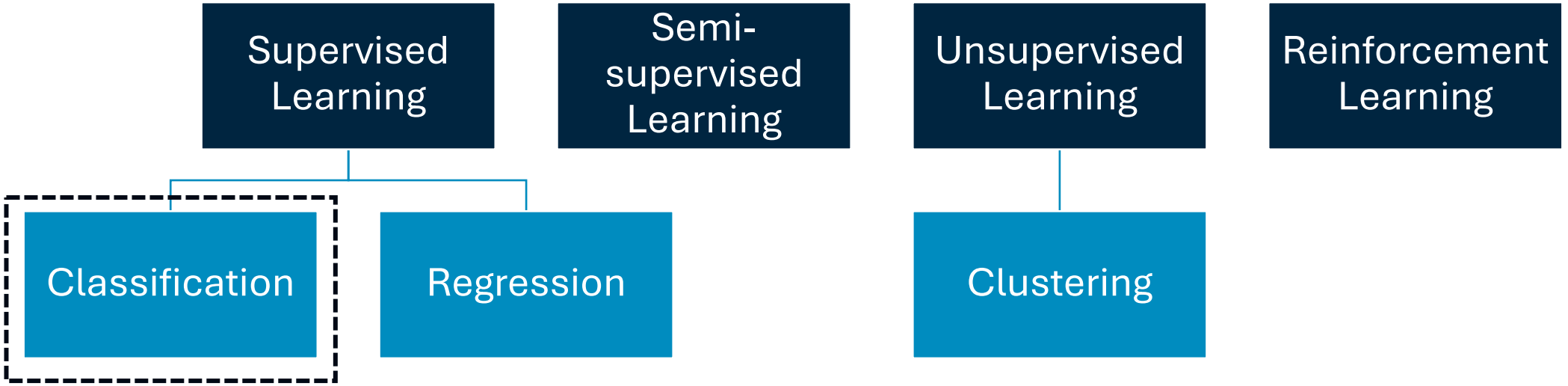
Sources: Jaime S. Cardoso (2018)

Now, we learn the rules directly from the data

We employ a **data-driven design**: the system produces a program that implements a **function** that assigns the **prediction** to any **observation**

Sources: Jaime S. Cardoso (2018)

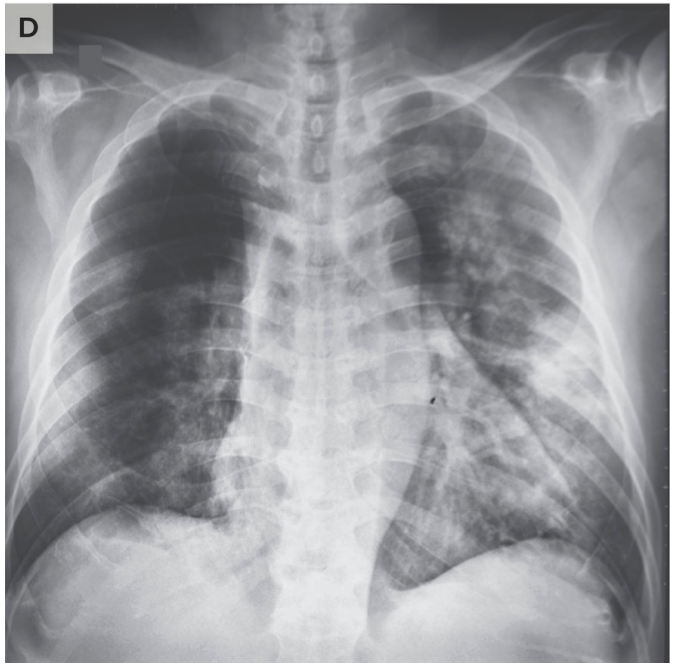
The taxonomy of machine learning is diverse: goals and available data dictate the type of learning problem



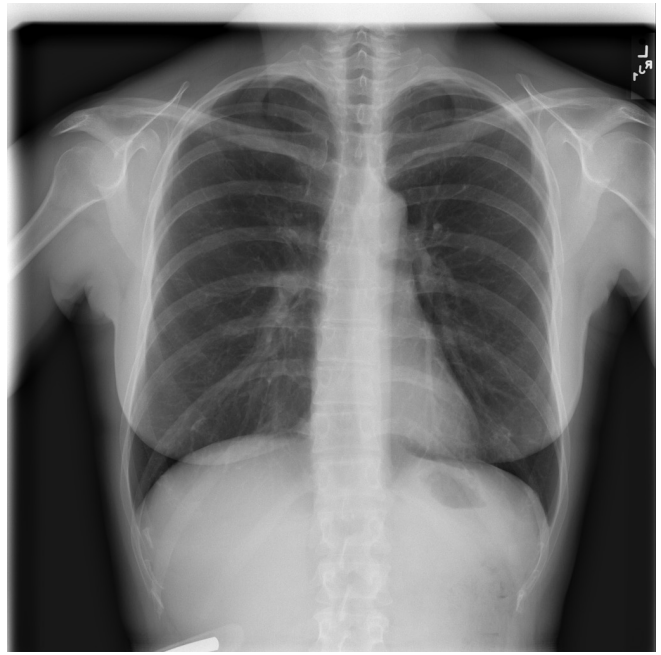
Sources: Jaime S. Cardoso (2018)

What if we had to build a computer-aided diagnosis system to detect COVID-19?

We would start by **understanding** the **data**[1]:



COVID-19

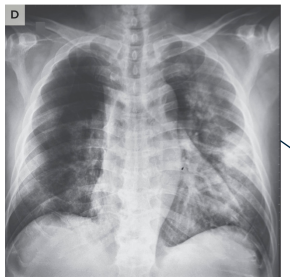


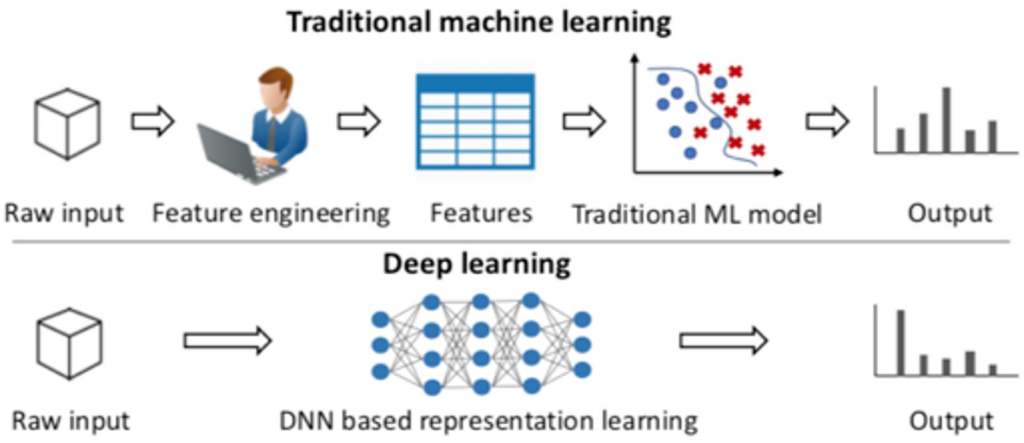
Normal

Sources: [1] Kaggle

What if we had to build a computer-aided diagnosis system to detect COVID-19?

Afterwards, we would **train** an **algorithm**:





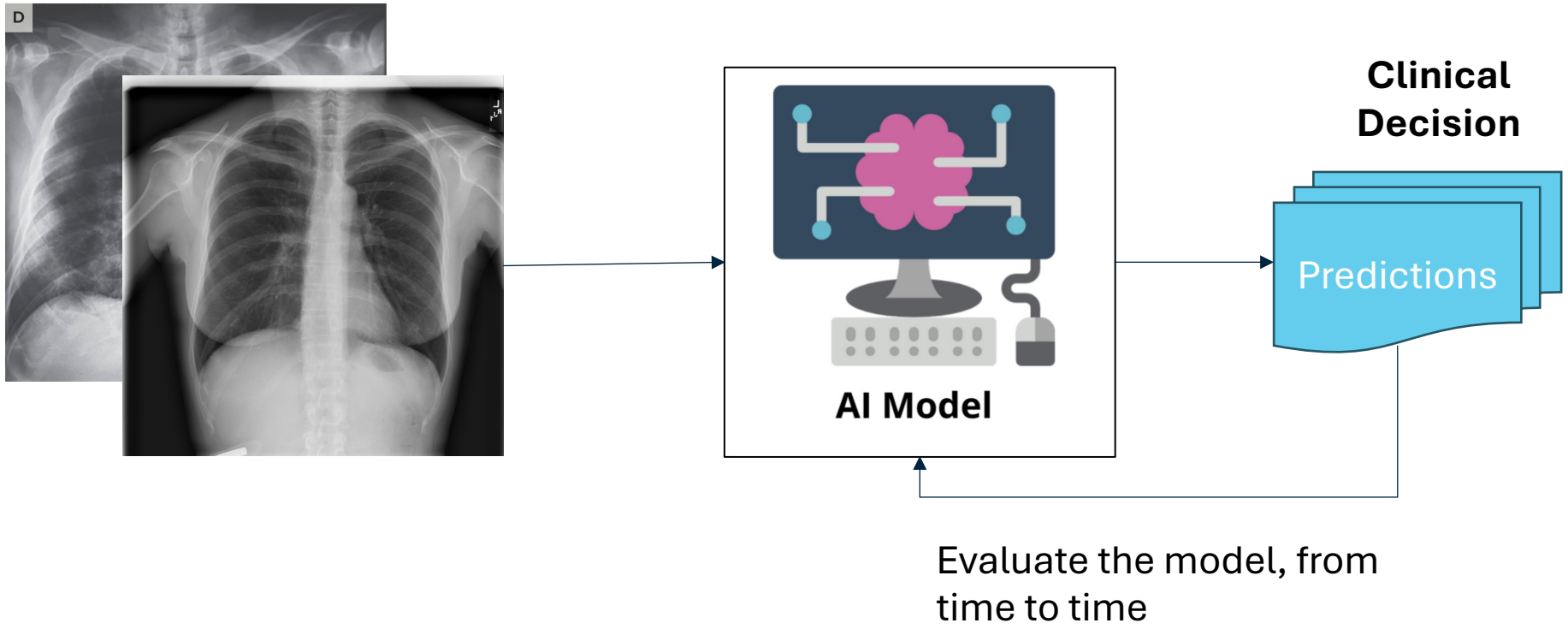
COVID-19 or Normal?

Optimise an objective function
(usually, minimising an error function)

Sources: Gunning (2019)

What if we had to build a computer-aided diagnosis system to detect COVID-19?

Finally, we would **deploy** the **model**:

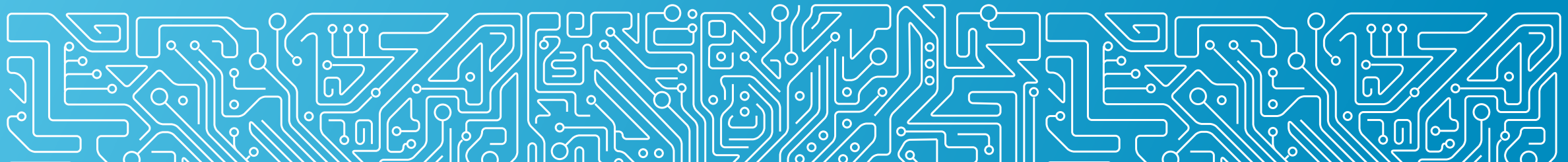


Sources: INESC TEC

INESC TEC

Epilogue

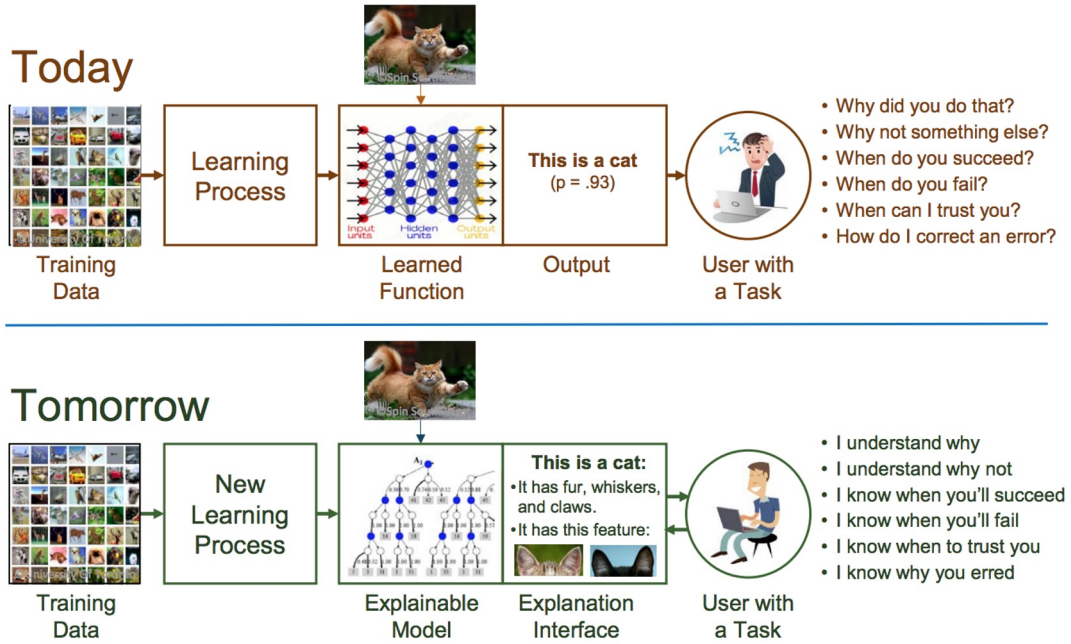
Future trends and open questions



Can we really understand what models are learning?

- It is not trivial to assure that models are **learning features** that are **relevant** for that **domain** (i.e., **black box behaviour**)

- Remember: **machine learning** models are **good** at extracting **correlations**!



Sources: Gunning (2019)

The Responsible AI framework is trying to bring more understandability to algorithms

Responsible AI is a framework that guides how we should address the challenges around AI from both an **ethical**, **technical** and **legal** point of view[1]

We must **resolve ambiguity** for where **responsibility** lies if something goes **wrong**!

This framework relies on **fundamental principles**[2]:

- Accountability
- Interpretability
- Fairness
- Safety
- Privacy

Sources: [1] Tech Target, [2] Towards Data Science

INESC TEC

INSTITUTO DE ENGENHARIA
DE SISTEMAS E COMPUTADORES,
TECNOLOGIA E CIÊNCIA

Learning How Machines Learn: An Introduction to Artificial Intelligence and Machine Learning

Tertúlia | Universidade Católica Portuguesa – Católica Porto Business School

July 7, 2025 | Porto, Portugal

Tiago Filipe Sousa Gonçalves (tiago.f.goncalves@inesctec.pt)