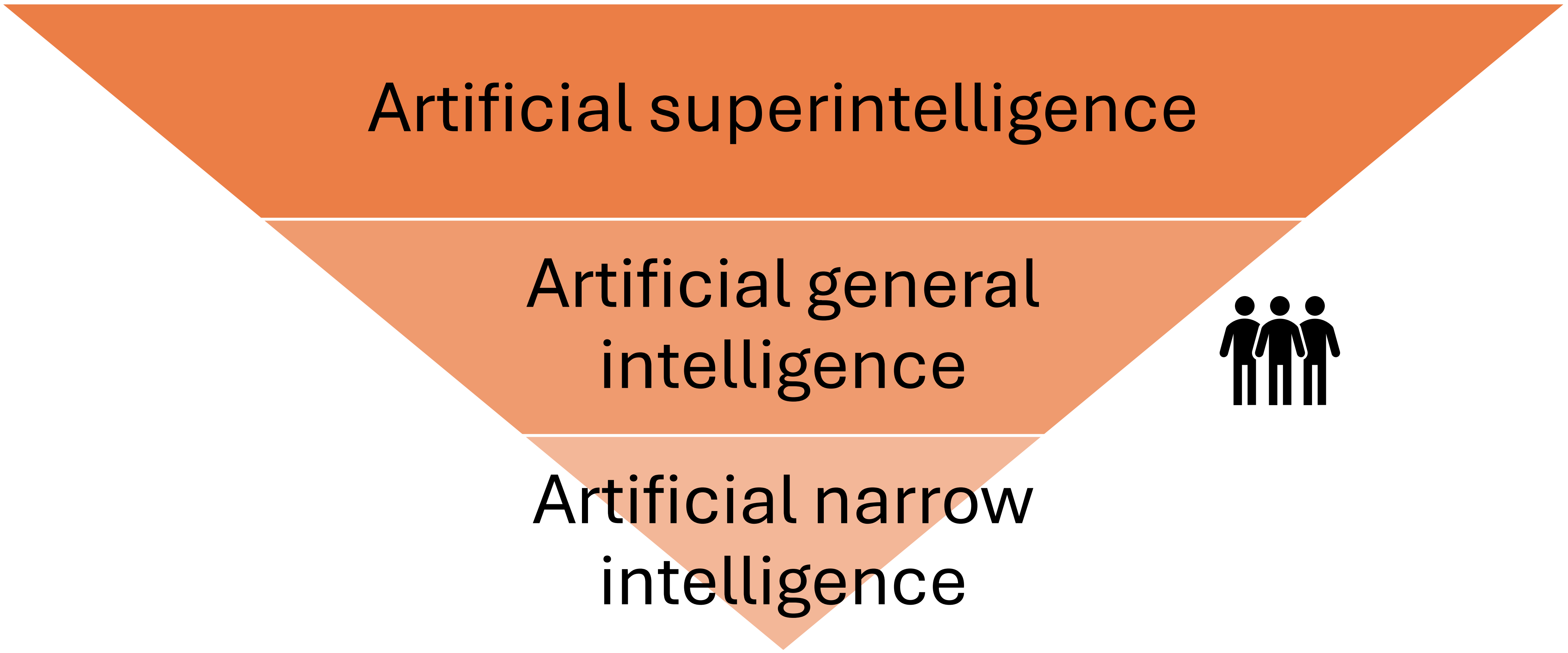
AI technology – part 1

Sonja Aits

Lund University

2025-09-22


Artificial superintelligence
Artificial general intelligence
Artificial narrow intelligence

Sparks of Artificial General Intelligence: Early experiments with GPT-4

Sébastien Bubeck Varun Chandrasekaran Ronen Eldan Johannes Gehrke
Eric Horvitz Ece Kamar Peter Lee Yin Tat Lee Yuanzhi Li Scott Lundberg
Harsha Nori Hamid Palangi Marco Tulio Ribeiro Yi Zhang

Microsoft Research

Abstract

Artificial intelligence (AI) researchers have been developing and refining large language models (LLMs) that exhibit remarkable capabilities across a variety of domains and tasks, challenging our understanding of learning and cognition. The latest model developed by OpenAI, GPT-4 [Ope23], was trained using an unprecedented scale of compute and data. In this paper, we report on our investigation of an early version of GPT-4, when it was still in active development by OpenAI. We contend that (this early version of) GPT-4 is part of a new cohort of LLMs (along with ChatGPT and Google's PaLM for example) that exhibit more general intelligence than previous AI models. We discuss the rising capabilities and implications of these models. We demonstrate that, beyond its mastery of language, GPT-4 can solve novel and difficult tasks that span mathematics, coding, vision, medicine, law, psychology and more, without needing any special prompting. Moreover, in all of these tasks, GPT-4's performance is strikingly close to human-level performance, and often vastly surpasses prior models such as ChatGPT. Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system. In our exploration of GPT-4, we put special emphasis on discovering its limitations, and we discuss the challenges ahead for advancing towards deeper and more comprehensive versions of AGI, including the possible need for pursuing a new paradigm that moves beyond next-word prediction. We conclude with reflections on societal influences of the recent technological leap and future research directions.

Predicting harvest yield for individual fields

Grouping pollen by shape

Deciding whether an ER patient is infectious

Reducing data dimensionality by removal of redundant features

Recommending physiotherapy exercises

Predicting body weight after therapy with GLP-1 agonist

Simulating scenarios for biodiversity loss

Identifying a butterfly species

Summarizing field notes

Prioritizing emergency calls

Producing artificial protein sequences

Identifying subgroups of cells with similar appearance

Predicting harvest yield for individual fields

Grouping pollen by shape

Deciding whether an ER patient is infectious

Reducing data dimensionality by removal of redundant features

Recommending physiotherapy exercises

Predicting body weight after therapy with GLP-1 agonist

Simulating scenarios for biodiversity loss

Identifying a butterfly species

Summarizing field notes

Prioritizing emergency calls

Producing artificial protein sequences

Identifying subgroups of cells with similar appearance

Predicting harvest yield for individual fields

Predicting body weight after therapy with GLP-1 agonist

Regression

= Predicting a numerical value

Grouping pollen by shape

Clustering

= Grouping by similiarity

Identifying subgroups of cells with similar appearance

Deciding whether an ER patient is infectious

Classification

= Assigning to pre-defined groups

Identifying a butterfly species

Reducing data dimensionality by removal of redundant features

Simplification

= Removing uninformative data

Summarizing field notes

Data generation

= Producing new data based on existing data

Simulating scenarios for biodiversity loss

Producing artificial protein sequences

Reasoning

= Making decisions or drawing conclusions based on available information and logic

Recommending physiotherapy exercises

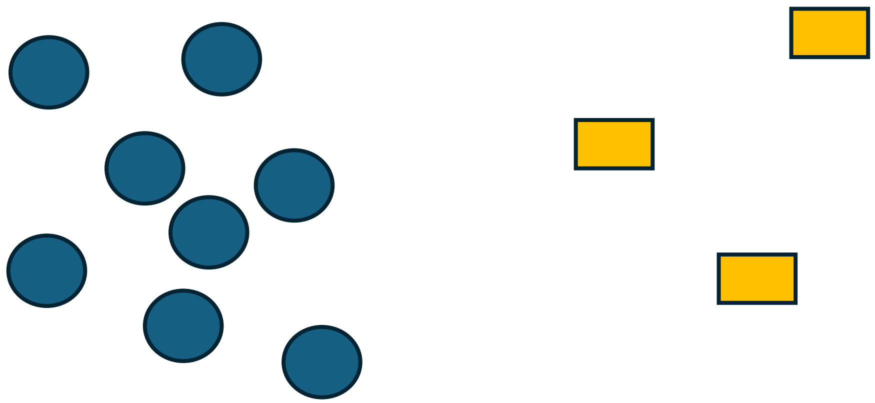
Prioritizing emergency calls

Find examples

1. Regression
2. Clustering
3. Classification
4. Simplification
5. Data generation
6. Reasoning

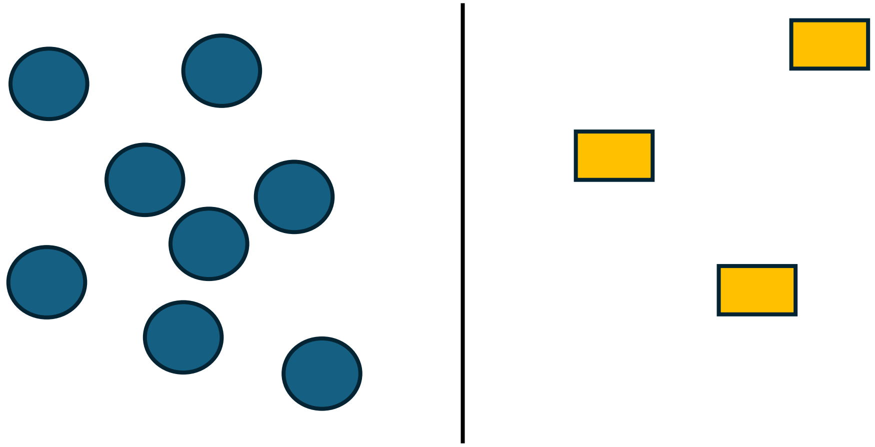
Many problems can be solved with different AI approaches

Finding faulty measurements
(Outlier detection)



Many problems can be solved with different AI approaches

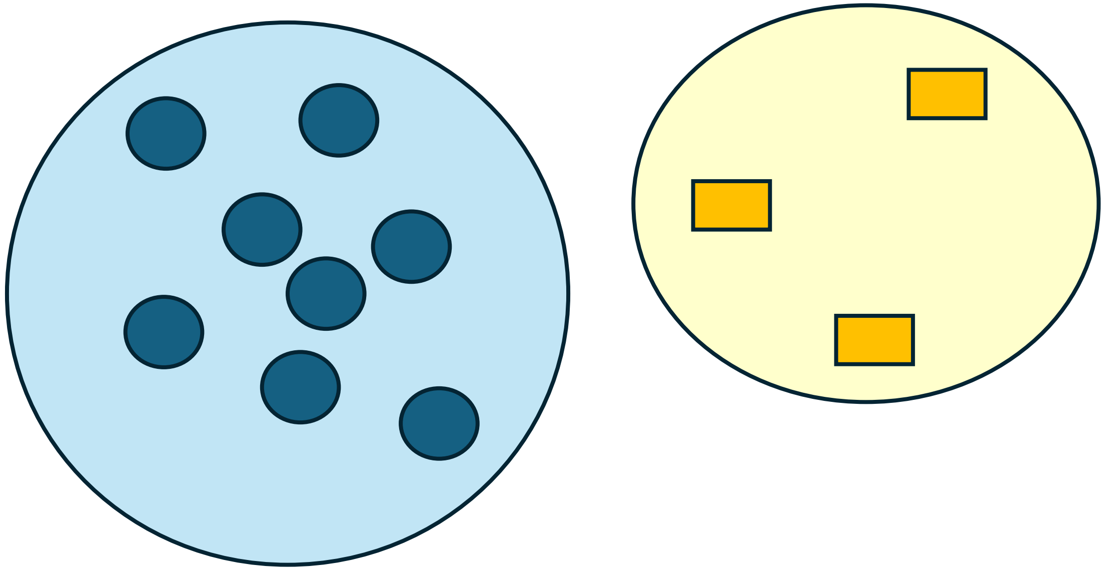
Finding faulty measurements
(Outlier detection)



Classification

Many problems can be solved with different AI approaches

Finding faulty measurements
(Outlier detection)



Clustering

Complex problems often require a combination of several AI core tasks

Targeted pesticide spraying

Classification: pest identification
Reasoning: pesticide selection, flight course plotting
Regression: amount of pesticide, rotor speed

AI comes in many shapes

Rule-based systems

Decision tree

K-nearest neighbor

K-means clustering

Linear regression

Convolutional neural network

Graph neural network

Transformer

Autoencoder

Recurrent neural network

Generative adversarial network

AI comes in many shapes

Rule-based systems

Machine learning

“Classical” machine learning

Decision tree

K-nearest neighbor

K-means clustering

Linear regression

Deep learning

Convolutional neural network

Graph neural network

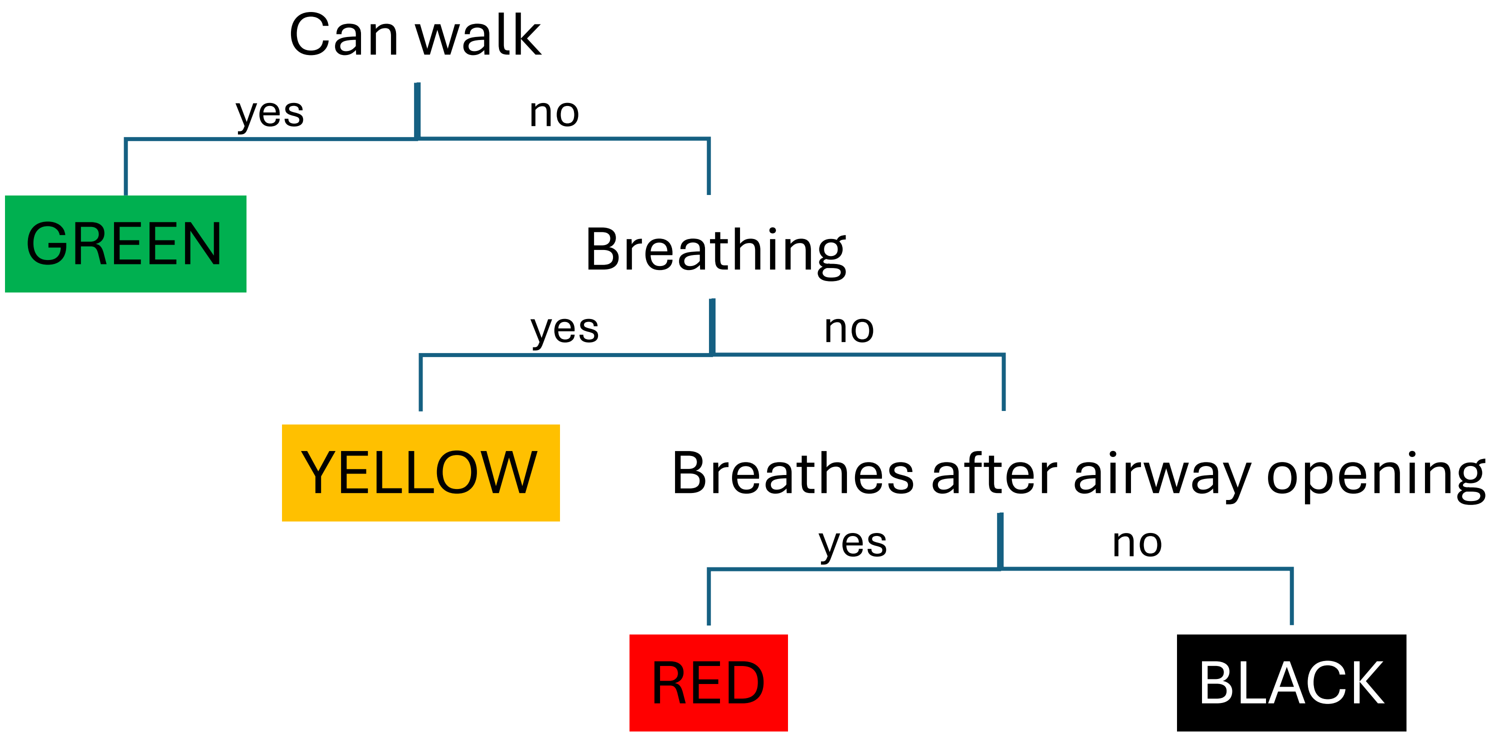
Transformer

Autoencoder

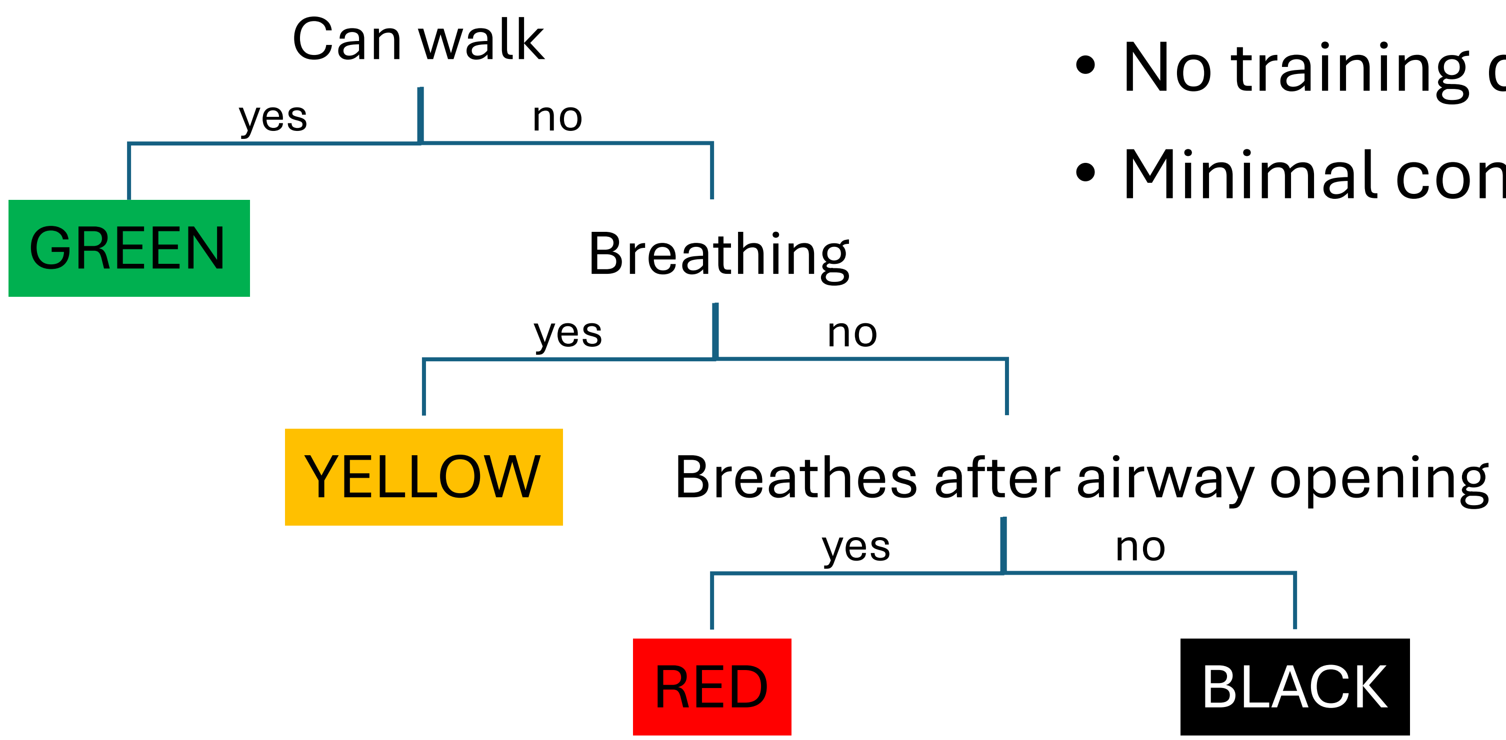
Recurrent neural network

Generative adversarial network

Rule-based systems encode human knowledge

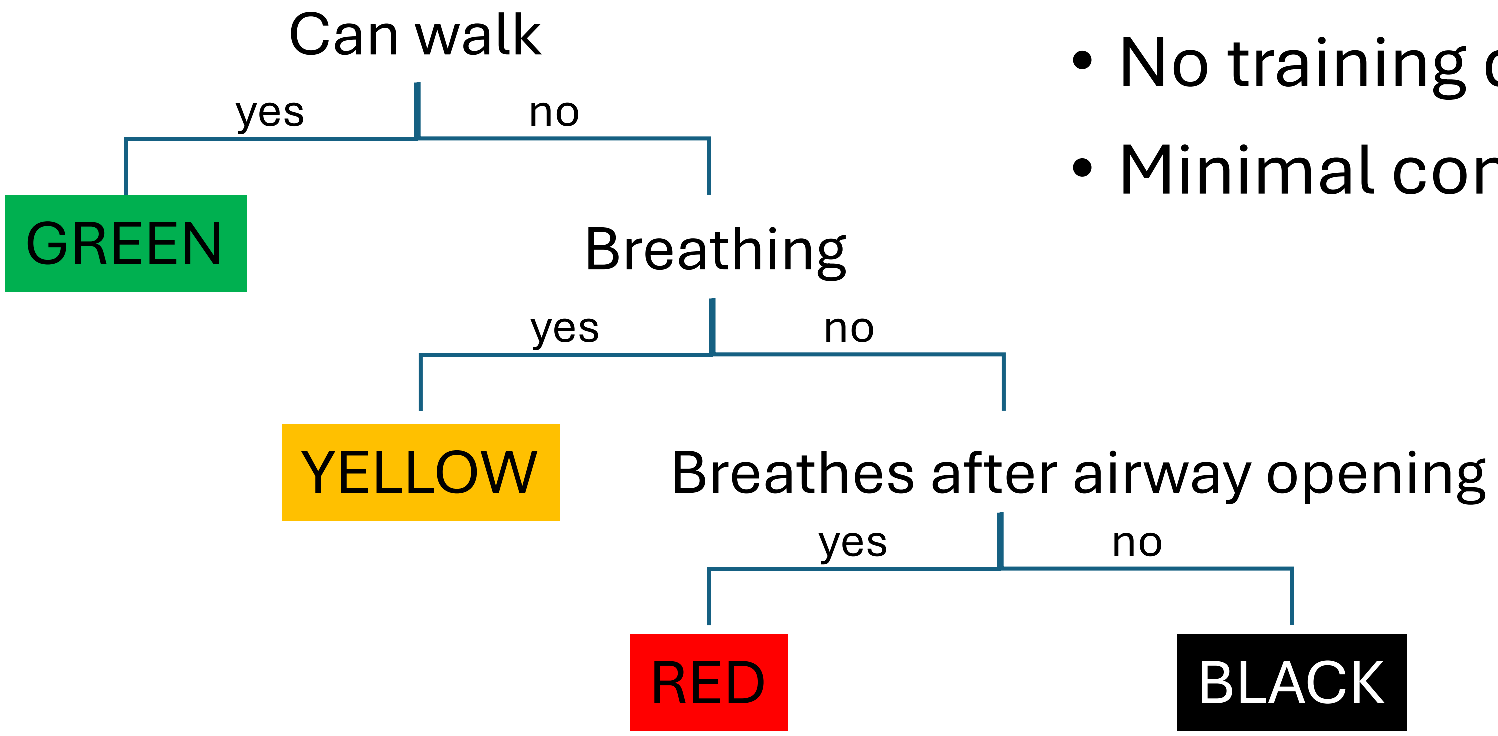


Rule-based systems encode human knowledge



- Fully explainable
- No training data required
- Minimal computational requirements

Rule-based systems encode human knowledge



- Fully explainable
- No training data required
- Minimal computational requirements

What if you cannot define the rules?

Machine learning

= learning from the data

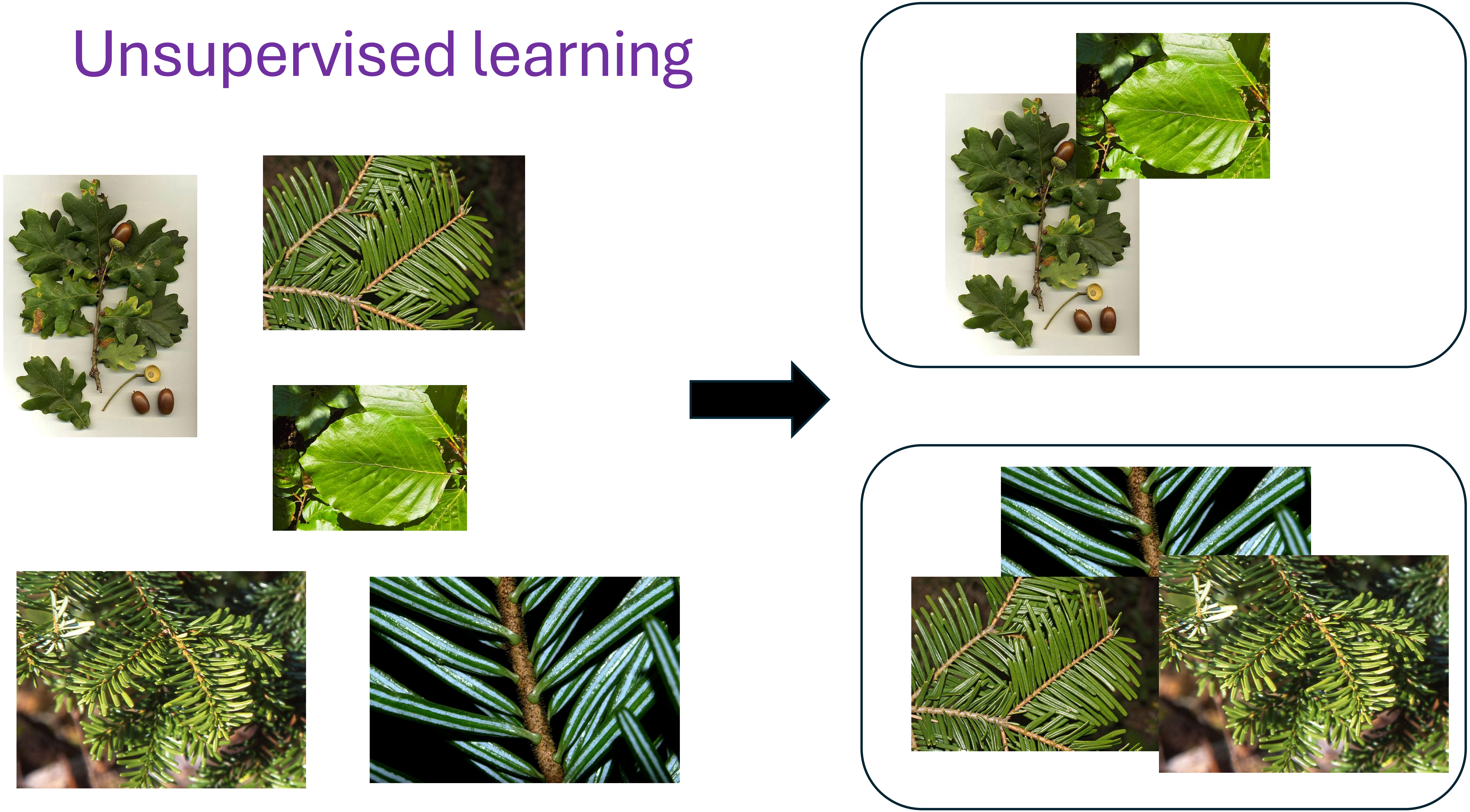
AI can be trained in multiple ways

Supervised learning	**Unsupervised learning**
Labeled data (x,y)	Unlabeled data
Defined target	No defined target
Aim: mapping function x -> y	Aim: Learn structure of the data
Common uses: regression, classification	Common uses: clustering, dimensionality reduction

Supervised learning

Oak

Unsupervised learning



Other ways to train AI

- **Semi-supervised learning**: learning from a mix of labeled and unlabeled data
- **Self-supervised learning**: learning from unlabeled data using pretext tasks for supervision
- **Reinforcement learning**: learning by rewarding/penalizing actions

Supervised or unsupervised?

- Predicting whether patient has COVID-19 from a CT scan of the lung
- Determining the severity grade of a tumor from histology images
- Clustering cells of similar shape from microscopy images
- Tracking bird flightpaths in a video
- Outlining areas of deforestation on satellite images
- Flagging blurry MRI images
- Finding unknown types of image corruption in a large dataset

Classical machine learning - examples

- **Decision trees**: like rule-based systems but algorithm creates the decision rules instead of human experts
- **K-nearest neighbor**: prediction based on k-most similar datapoints seen during training
- **K-means clustering**: data grouped into k clusters based on similarity
- **Linear regression**: models linear relationship between dependent variable and one or more independent variables ("fitting a line/hyperplane"

Classical machine learning – advantages compared to deep learning

- Higher explainability
- Often less data required
- Usually faster to train
- Require less computational resources
- Often superior performance with structured (tabular) data

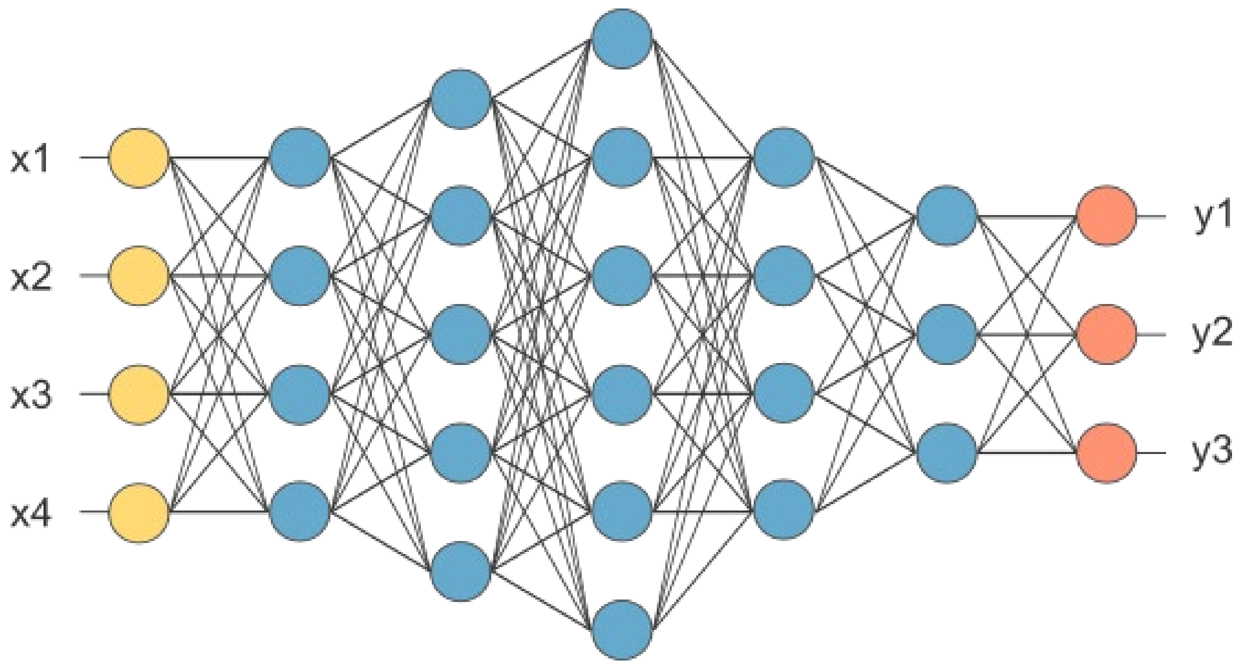
Classical machine learning – advantages compared to deep learning

- Higher explainability
- Often less data required
- Usually faster to train
- Require less computational resources
- Often superior performance with structured (tabular) data

BUT:
often insufficient for complex tasks with unstructured data
→ Deep learning

What are deep neural networks?

Deep neural networks are chains of interconnected mathematical operations



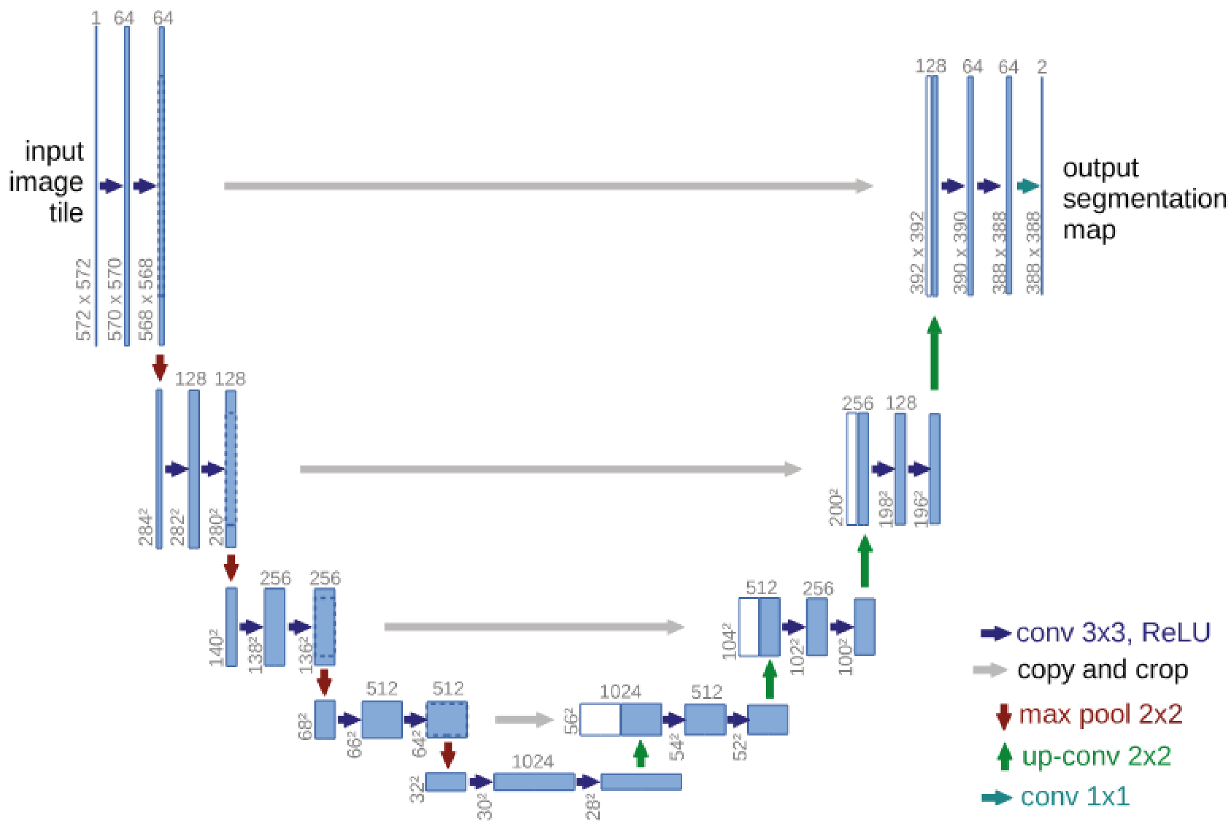
Input layer

Hidden layers

Output layer

Neural network shape and notation is highly variable

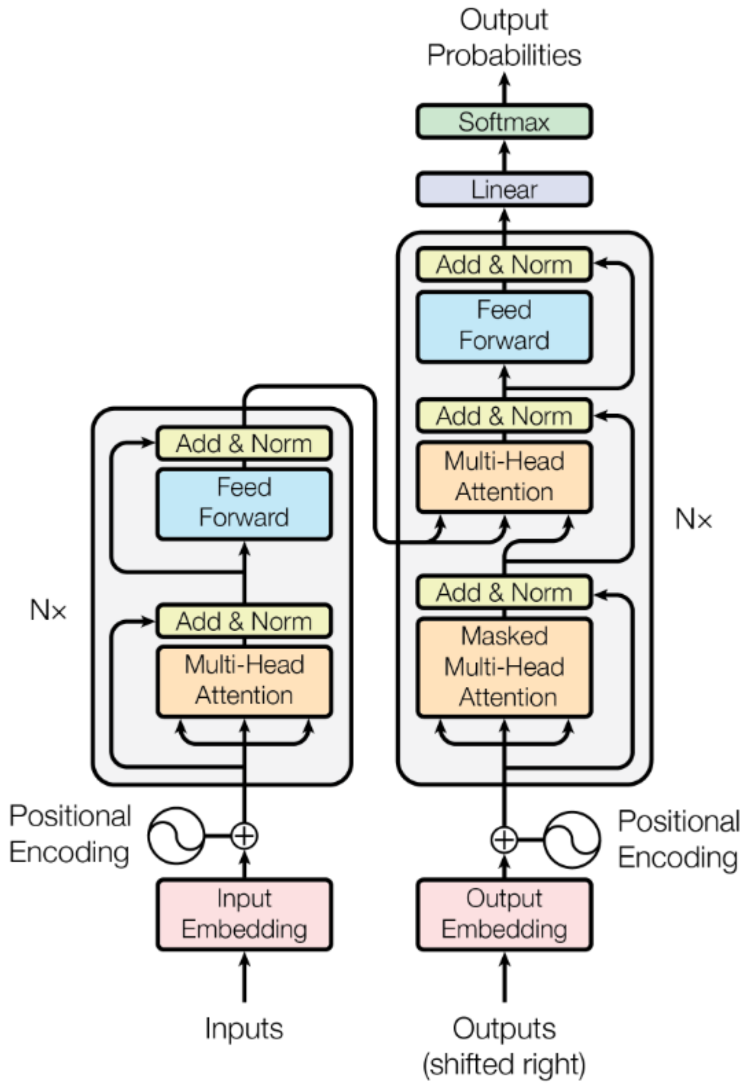
Input layer



Output layer

Hidden layers

Neural network shape and notation is highly variable

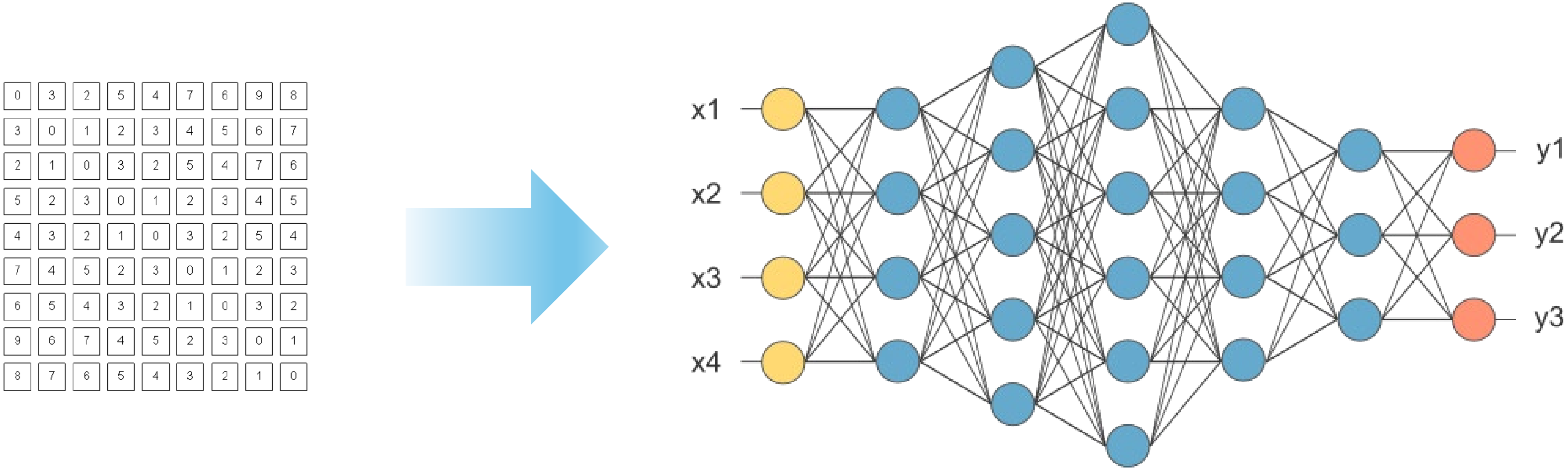


How we perceive the world

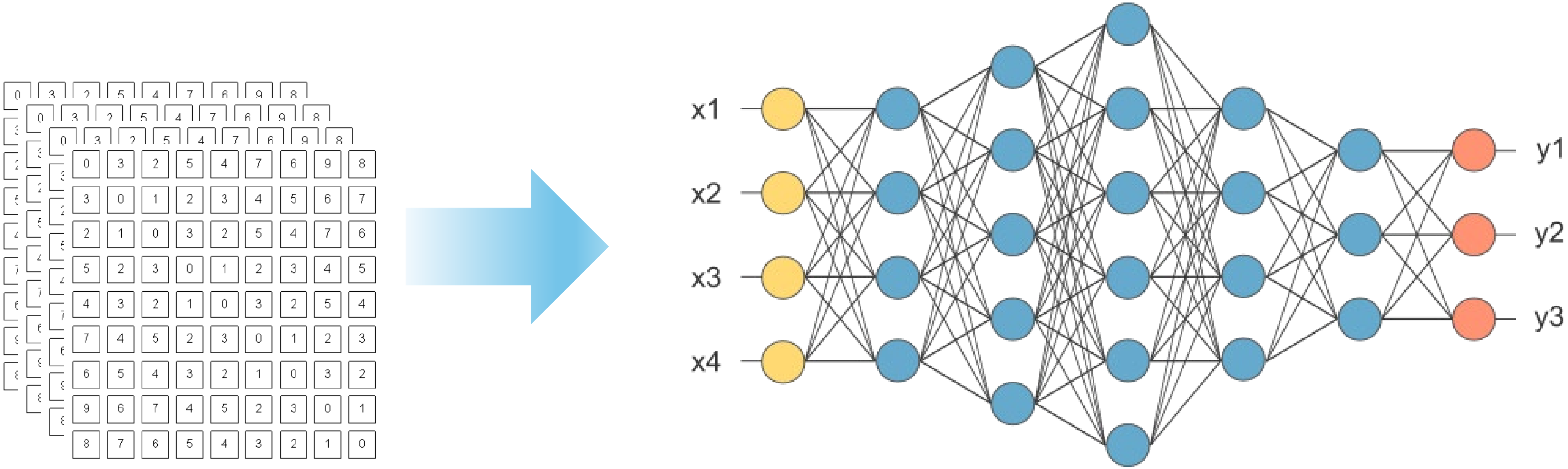
How computers perceive the world

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

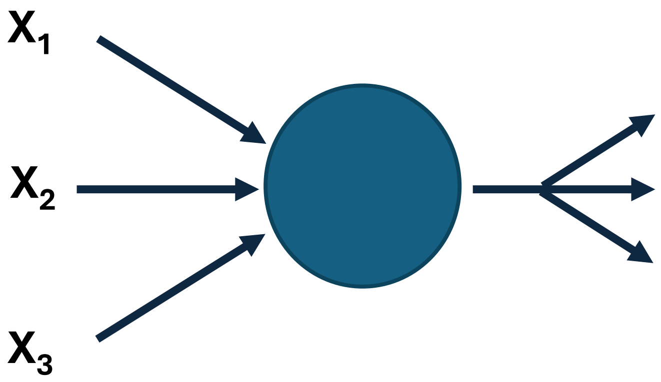
Deep neural networks are chains of interconnected mathematical operations

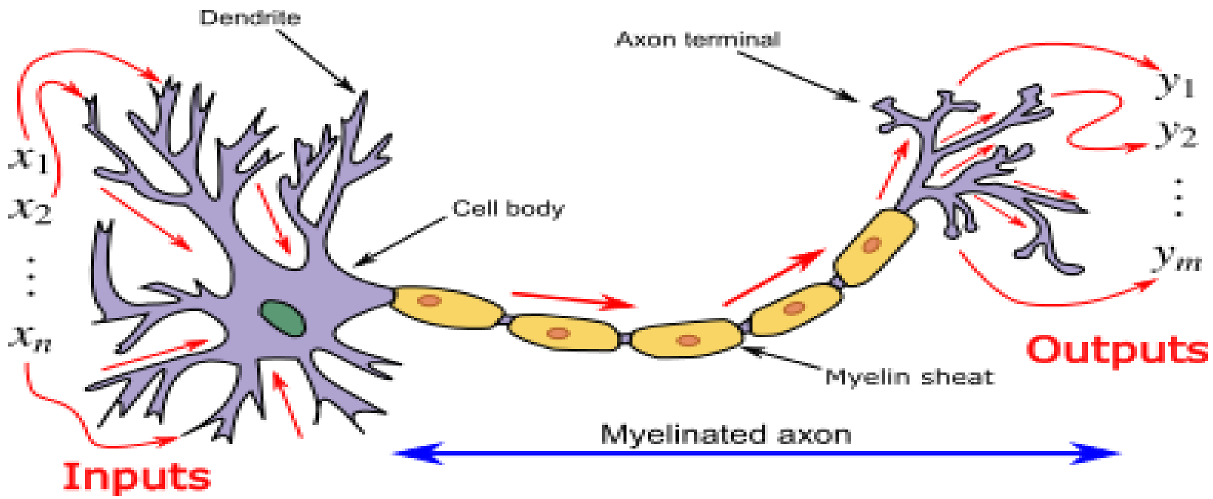


Deep neural networks are chains of interconnected mathematical operations

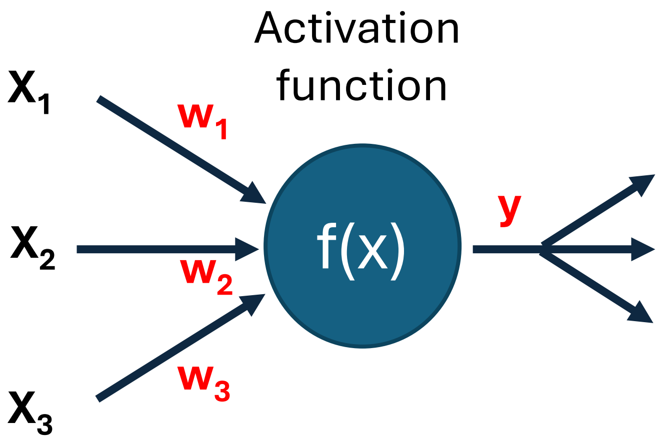


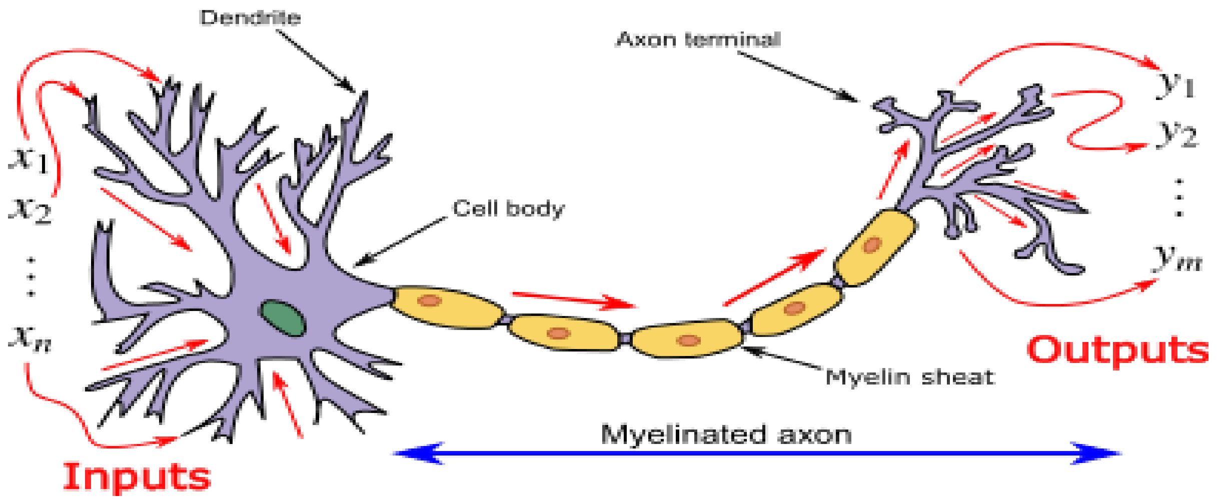
Each network node can have several input and output connections



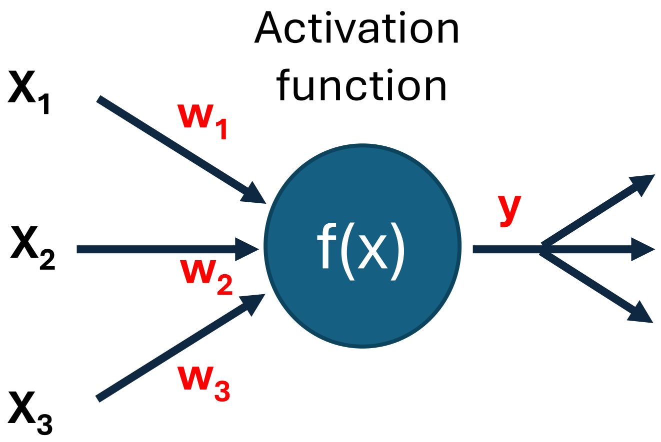


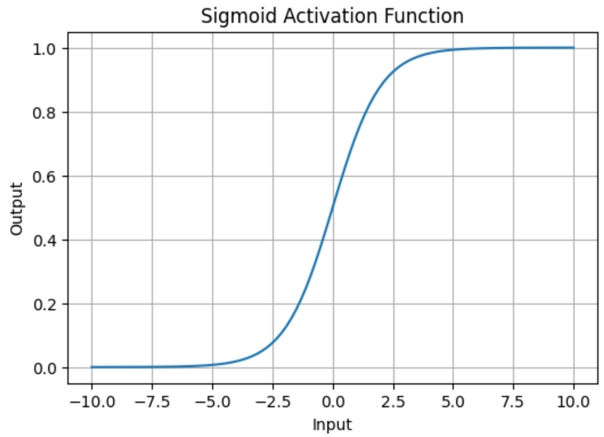
In each node inputs are weighted, aggregated and transformed to an output

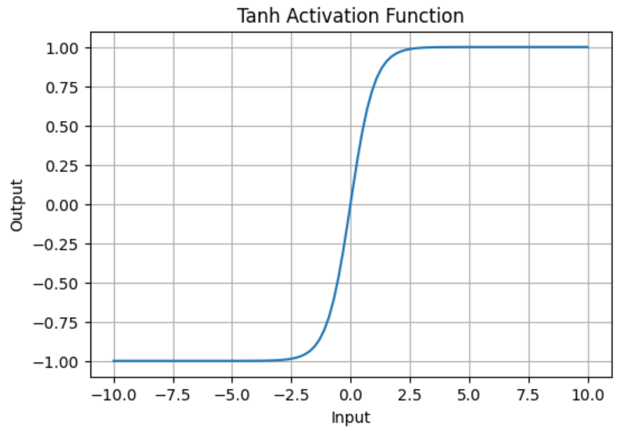


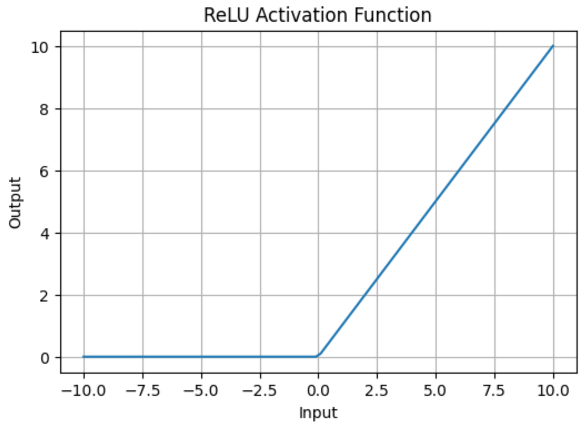


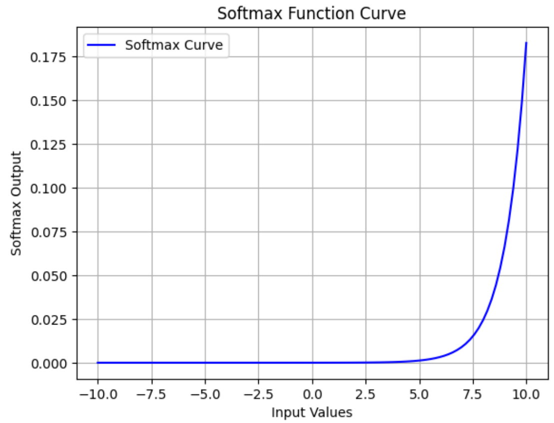
Activation functions are a crucial part of the network architecture





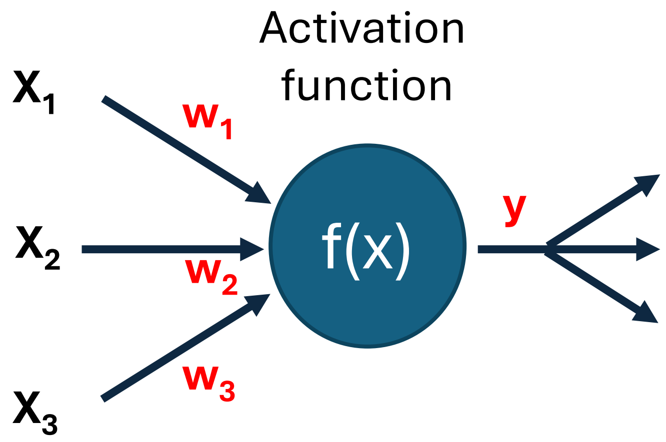


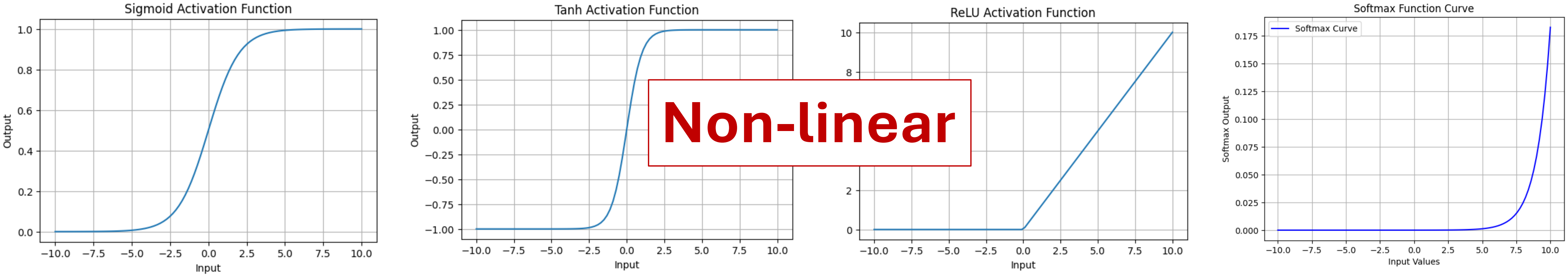




https://www.geeksforgeeks.org/machine-learning/activation-functions-neural-networks/

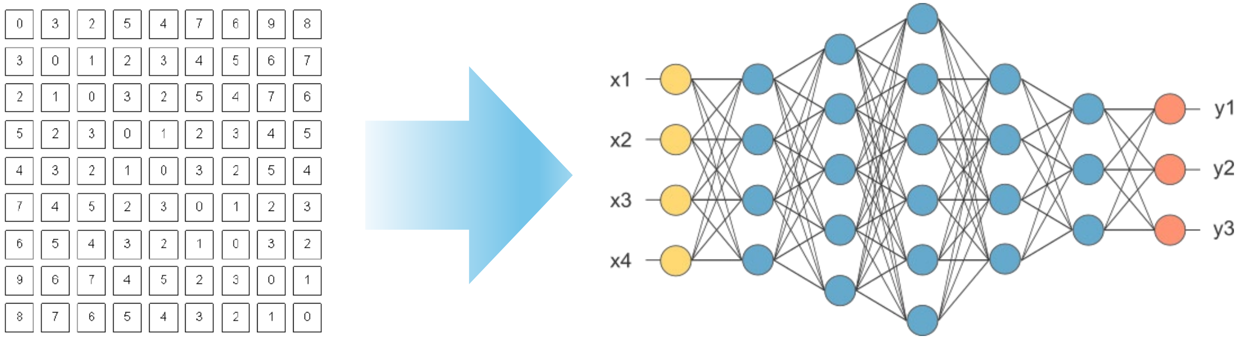
Activation functions are a crucial part of the network architecture





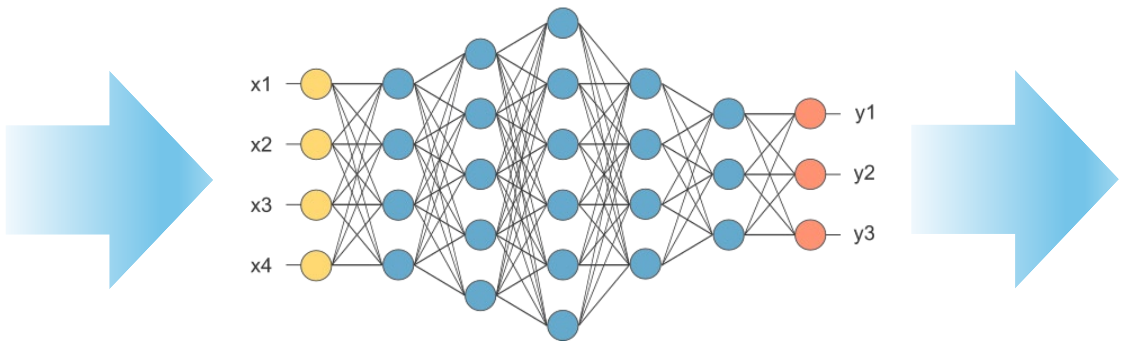
https://www.geeksforgeeks.org/machine-learning/activation-functions-neural-networks/

How to train a neural network...



How to train a neural network...

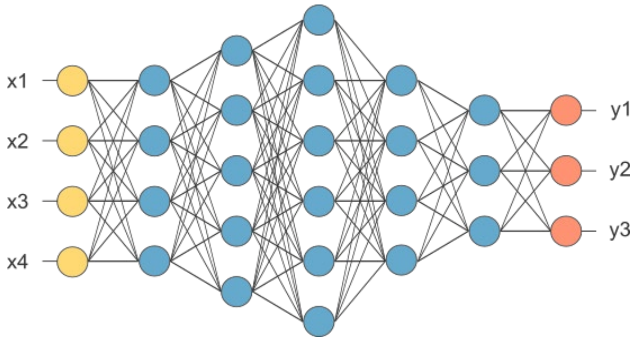
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0



Output

How to train a neural network…

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0



Output

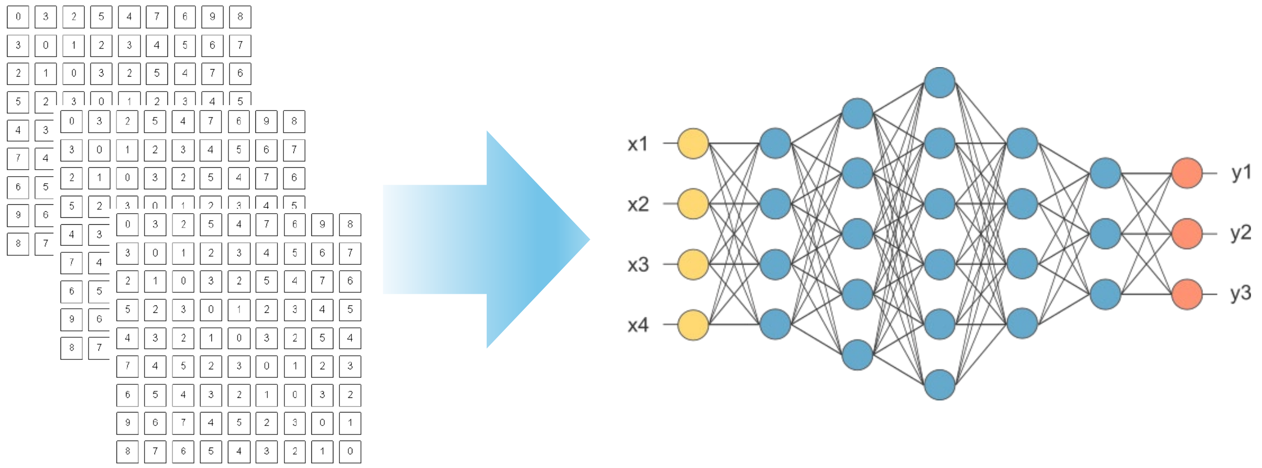
Error calculation (Loss function)

Loss function is chosen depending on the task

- Classification: Binary Cross-Entropy, Categorical Cross Entropy
- Regression: Mean Squared Error (MSE), Mean Absolute Error (MAE)
- …

Loss function = measure of success

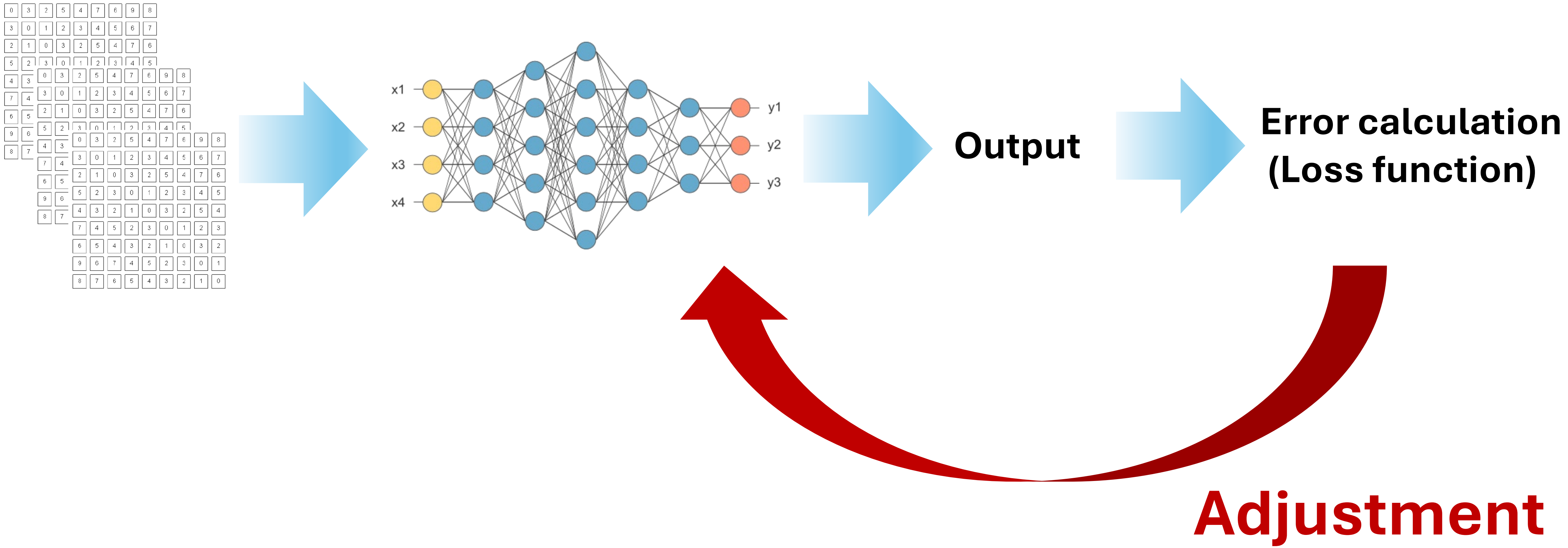
Error is calculated based on multiple input examples



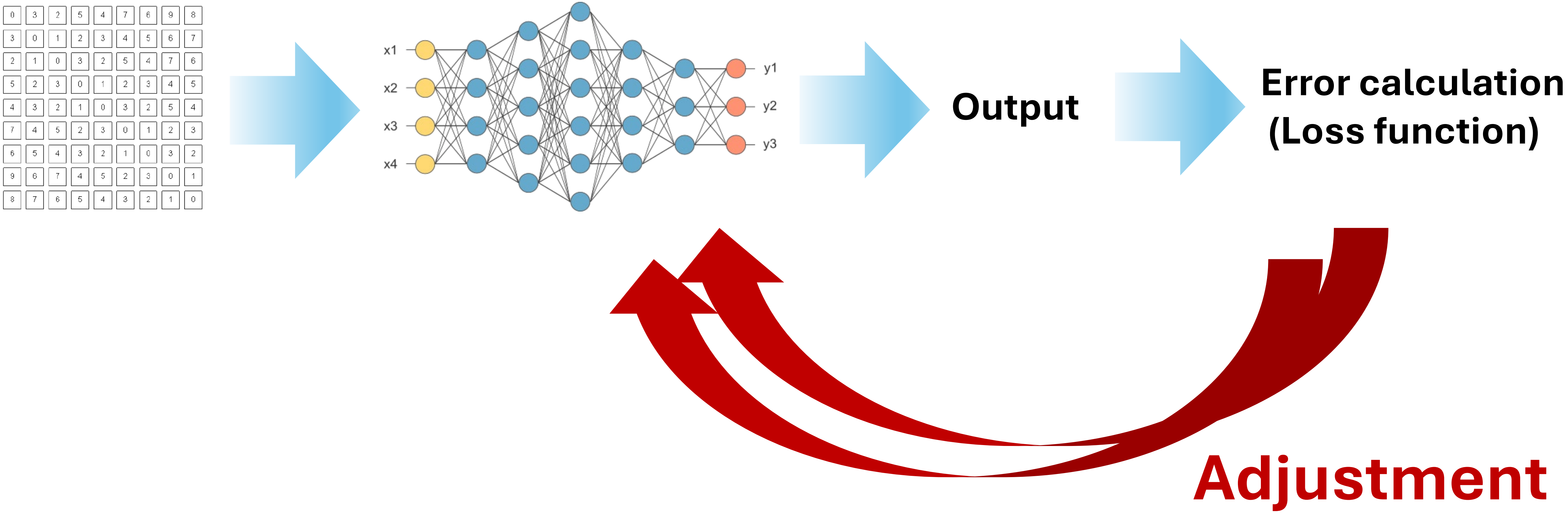
Output

Error calculation (Loss function)

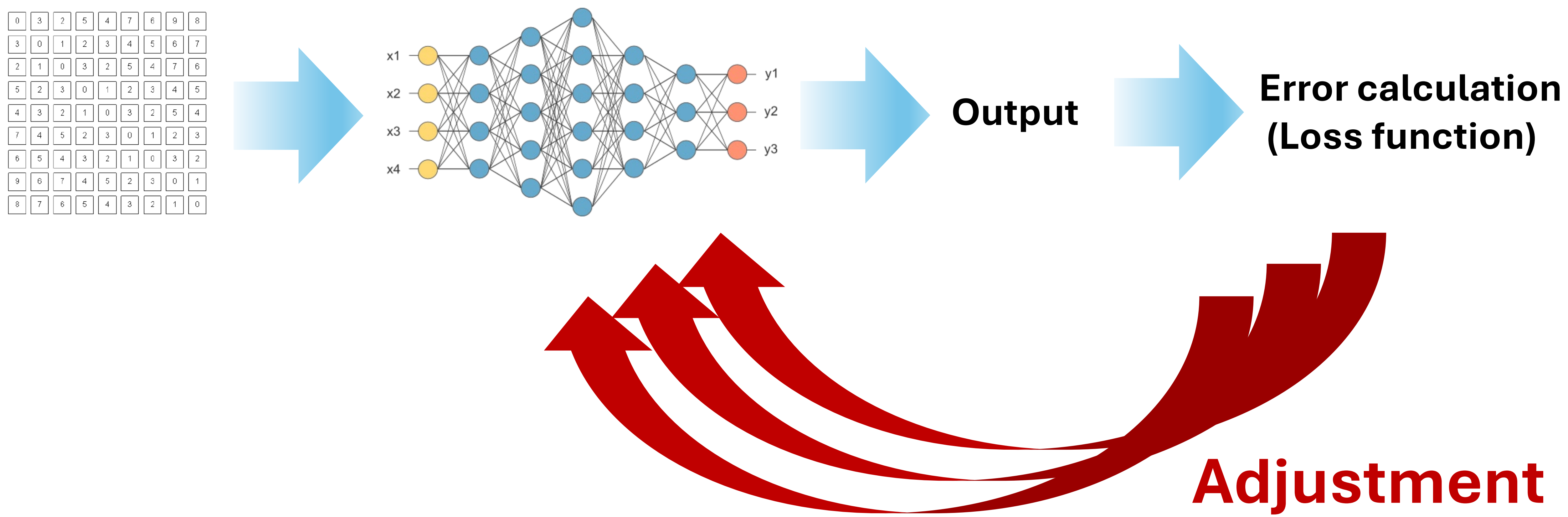
Weights and other parameters are adjusted to reduce errors



Weights and other parameters are repeatedly adjusted to reduce errors



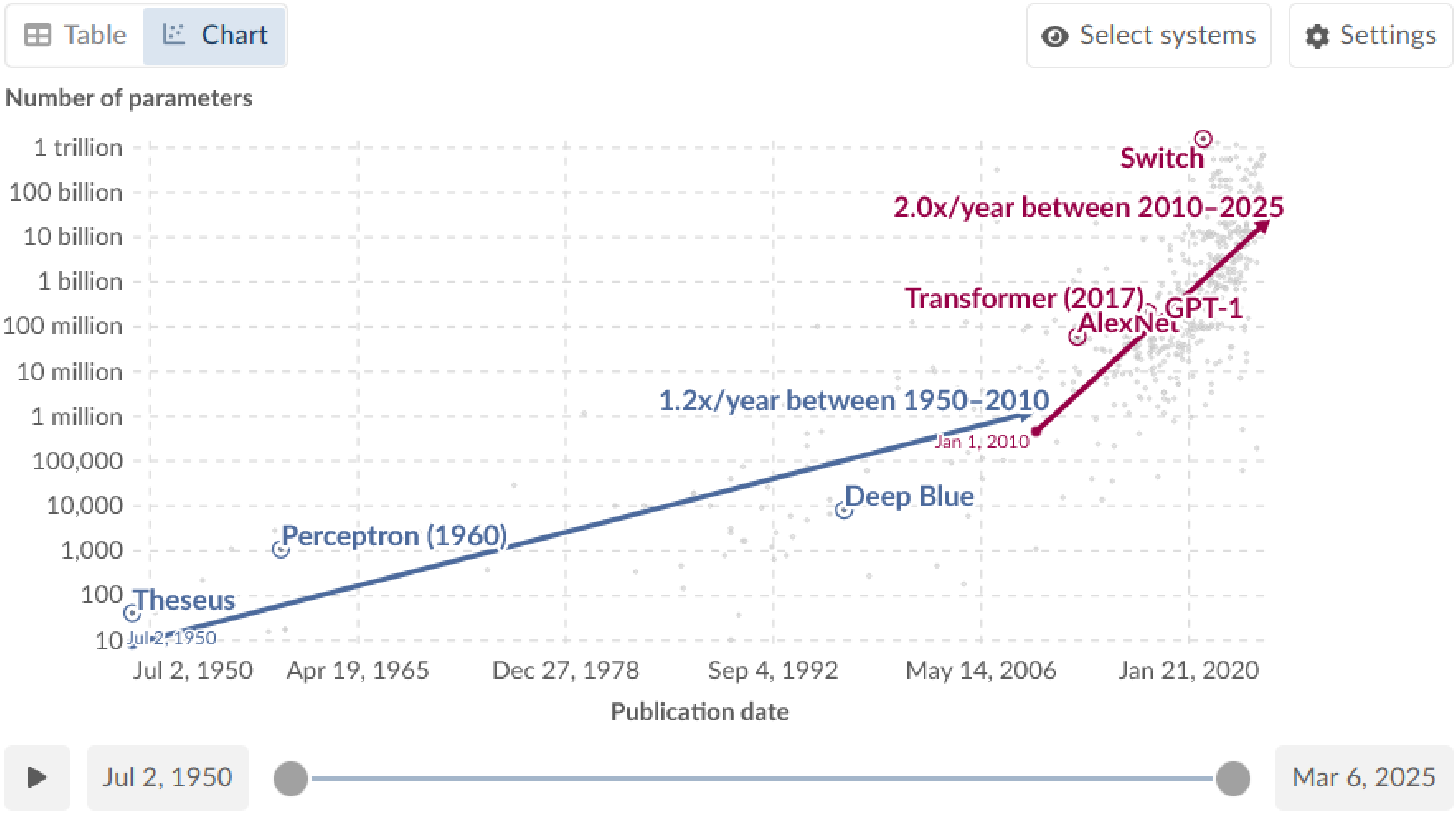
Weights and other parameters are repeatedly adjusted to reduce errors



Our World in Data

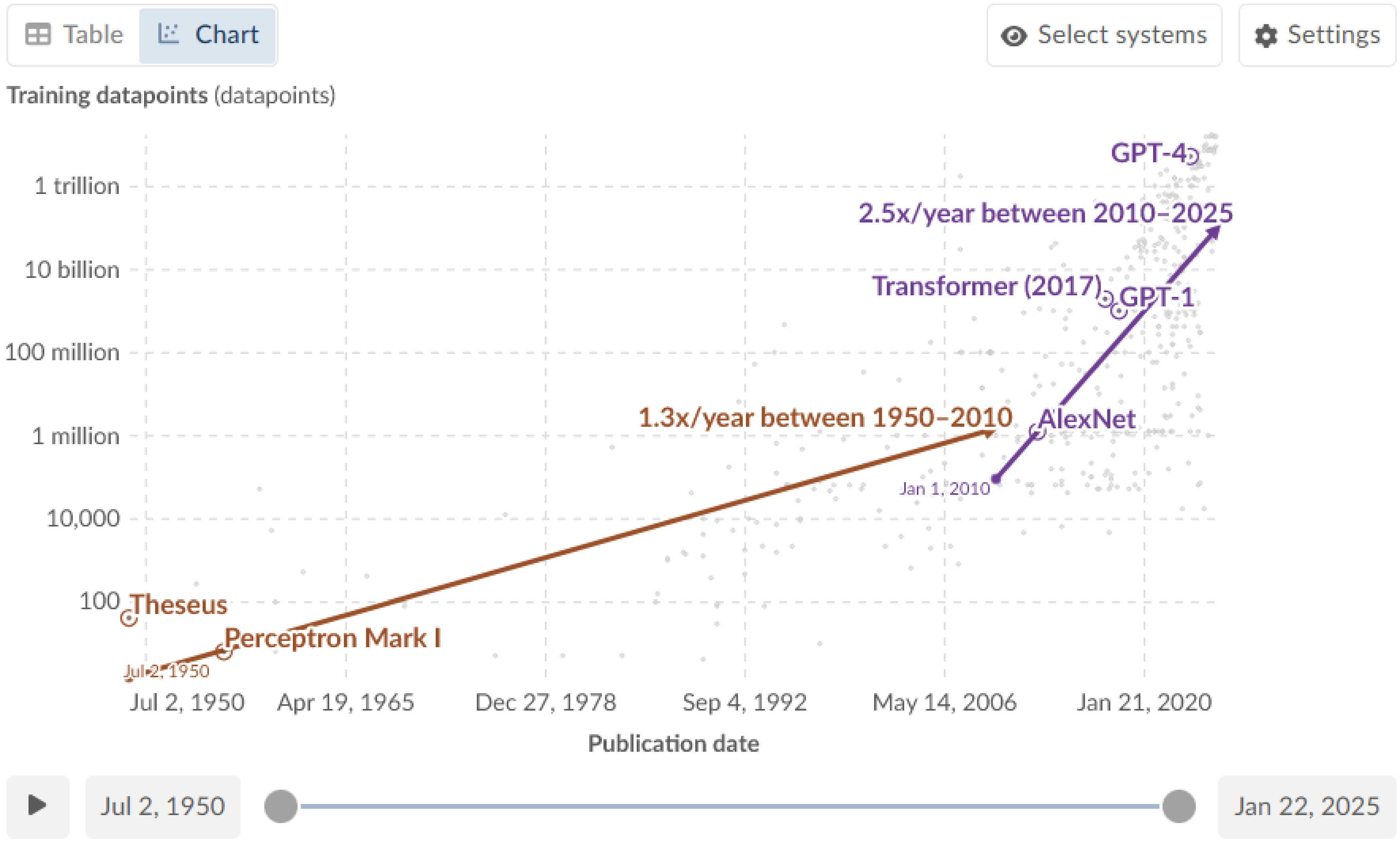
Exponential growth of parameters in notable AI systems

Parameters are variables in an AI system whose values are adjusted during training to establish how input data gets transformed into the desired output; for example, the connection weights in an artificial neural network.

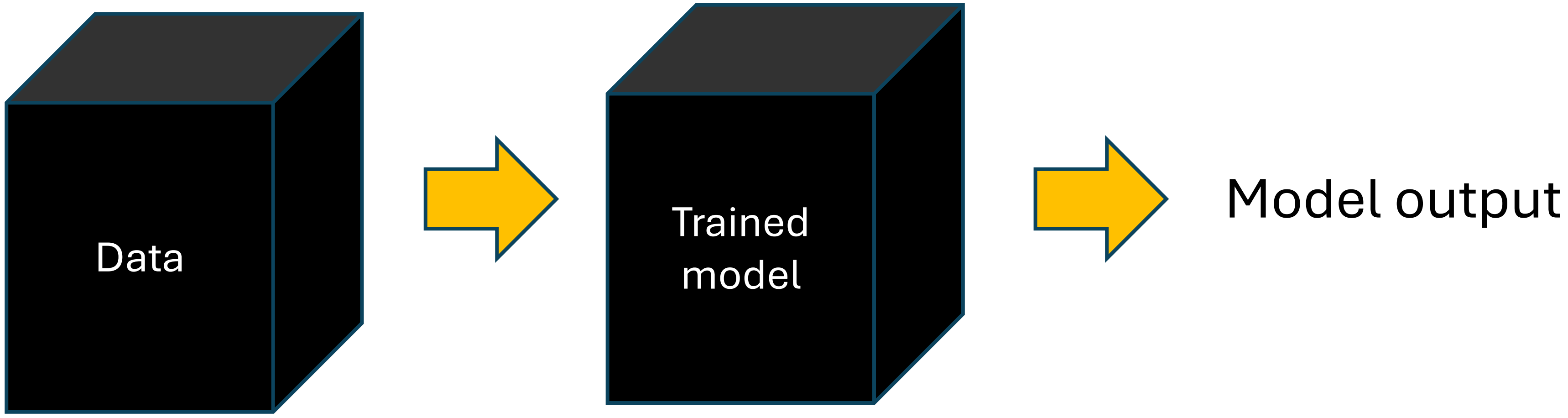


Exponential growth of datapoints used to train notable AI systems

Each domain has a specific data point unit; for example, for vision it is images, for language it is words, and for games it is timesteps. This means systems can only be compared directly within the same domain.



Deep learning models lack transparency “Black box issue”



Summary

- Current AI is “narrow” AI – mostly specialized for single task
- Problems can be broken down to core AI tasks (e.g. regression, classification)
- Most of today’s AI is based on machine learning - either classical machine learning or deep learning
- Deep learning = machine learning with deep neural networks
- Deep neural networks are trained by gradually adjusting parameters to minimize errors