# TOWARDS DATA DRIVEN AND MACHINE LEARNING BASED CONTROL

## Valmet Automation Seminar 2025

Prof. Eduard Petlenkov,
eduard.petlenkov@taltech.ee

Energy Efficiency Centre of Excellence
ENERGIATÕHUSUSE TIPPKESKUS

TAL TECH

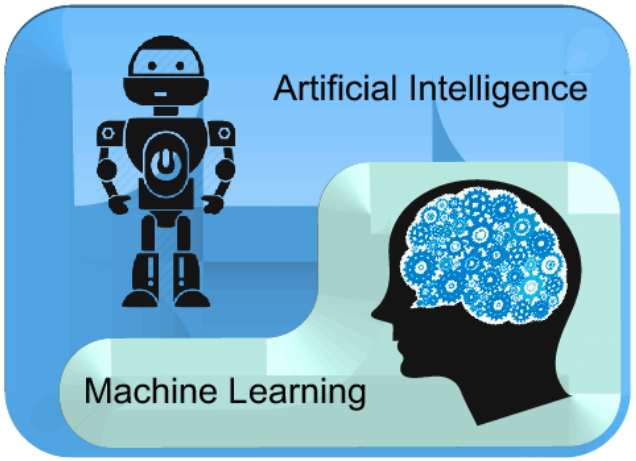
# ARTIFICIAL INTELLIGENCE VS MACHINE LEARNING

**AI** is a technology that enables a machine to simulate human behaviour (decision-making or problem solving).

The **focus** of AI is on solving problems.

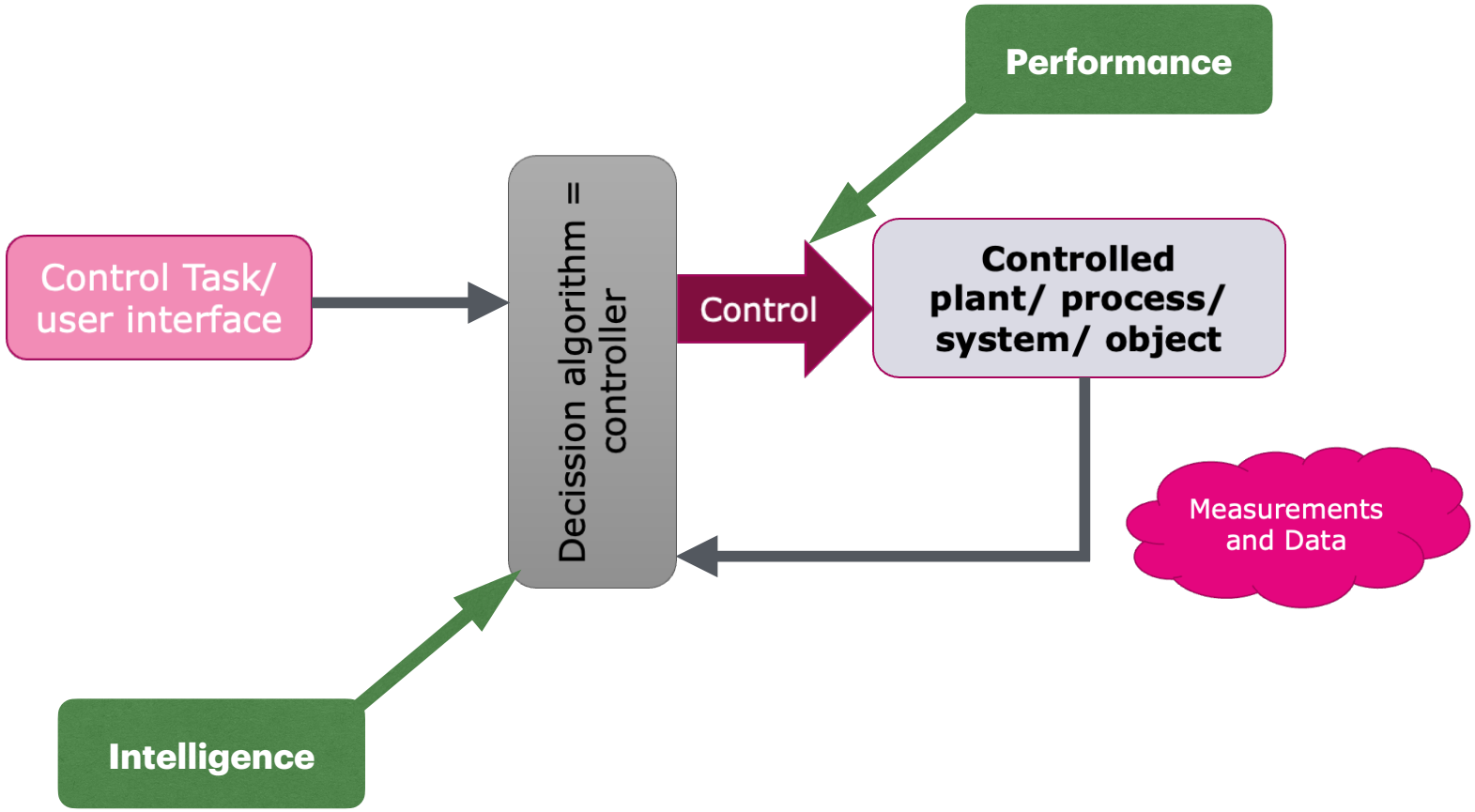
**ML** is a subset of AI which allows a machine to automatically learn from past data without programming explicitly.

The **focus** of ML is  on accuracy.

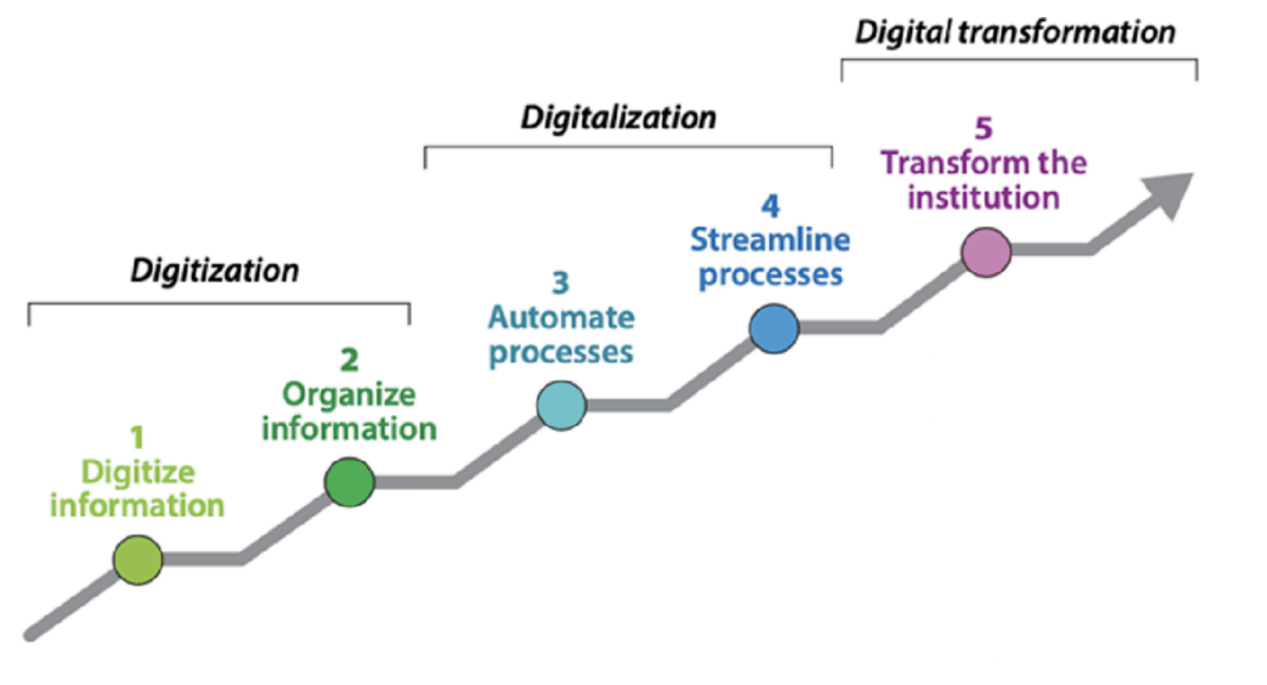



https://www.javatpoint.com/difference-between-artificial-intelligence-and-machine-learning

# INTELLIGENT CONTROL



# TRANSITION STEPS FROM NON-DIGITIZED WORLD TO THE DIGITIZED ONE



# MACHINE LEARNING

- Requires data
- Data has to be preprocessed/ presented in a suitable form
- Data should be reliable, clean, sufficient etc.

## Remember the main rule:

**Garbage in - Garbage out!**

# TRAINING/LEARNING

## SUPERVISED

Learning by example

Output is known

Suitable for classification and regression

## UNSUPERVISED

Learning by reasoning

Target is not provided

Suitable for clusterisation

## REINFORCEMENT

Learning by interacting with the environment

Target is provided

Output is unknown




https://arshren.medium.com/supervised-unsupervised-and-reinforcement-learning-245b59709f68

# SUPERVISED LEARNING

- Classification

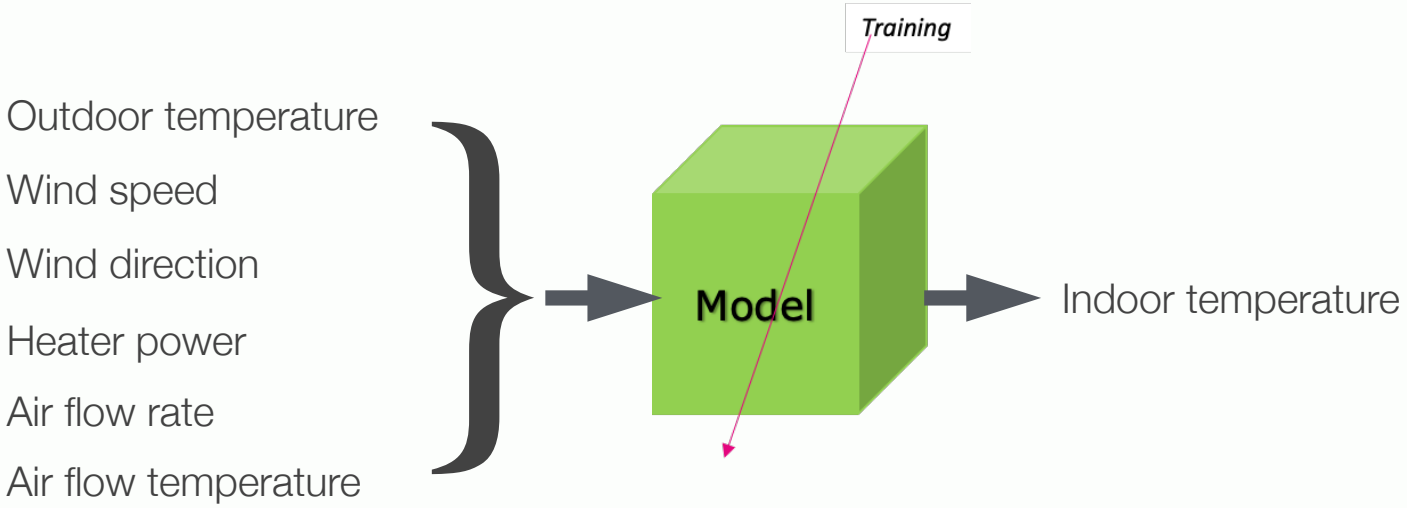
Assigning objects/inputs/patterns to some pre-existing classes or categories.

- Regression

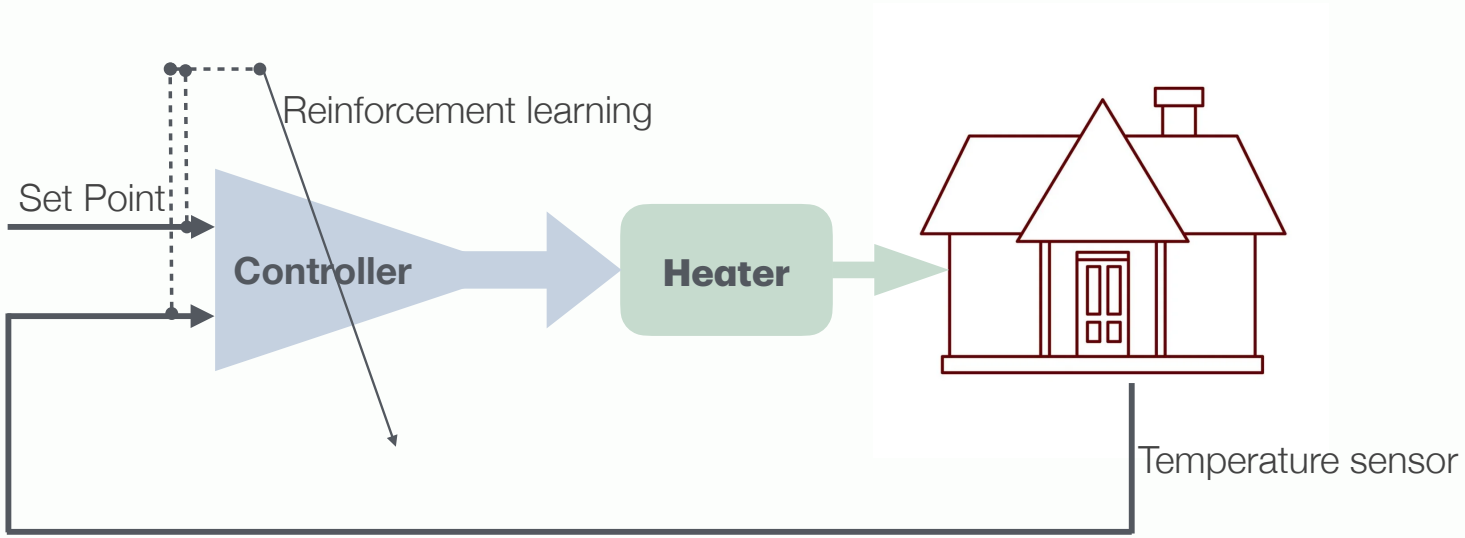
Determining the relationship between a dependent variable and one or more independent variables. Output is a real number.

# SUPERVISED LEARNING

## Regression Example



# REINFORCEMENT LEARNING

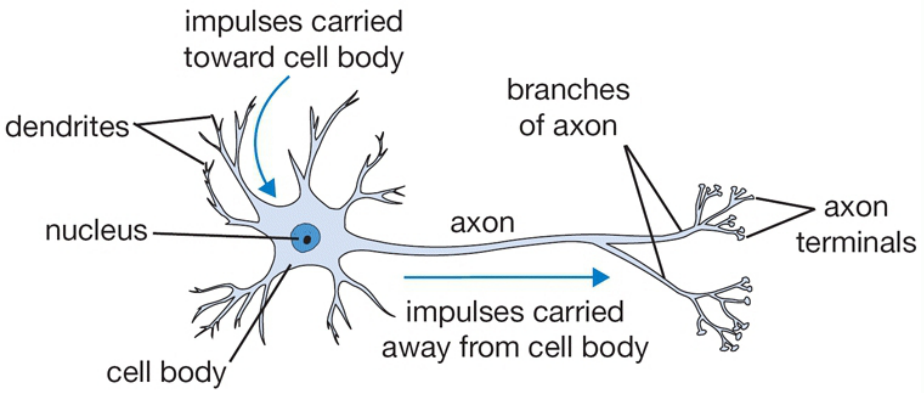


# Artificial Neural Networks

# STRUCTURE OF A BIOLOGICAL NEURON

A biological neuron has three types of main components:

**dendrites**, soma(or **cell body**) and **axon**.




Neurons pass information from one to another using action potentials. They connect with one another at **synapse**, which are junctions between one neuron's axon and another's dendrite.

Information flows from:

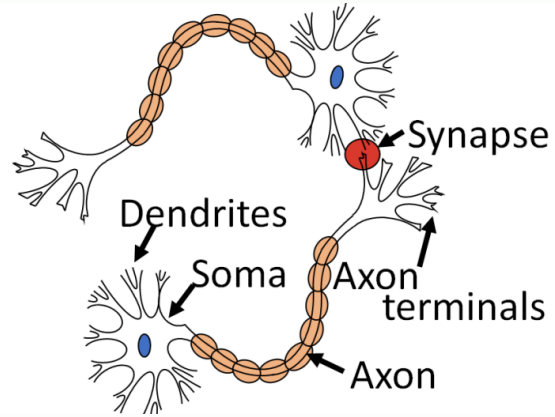
- the dendrite
- to the cell body
- through the axons
- to a synapse connecting the axon to the dendrite of the next neuron.

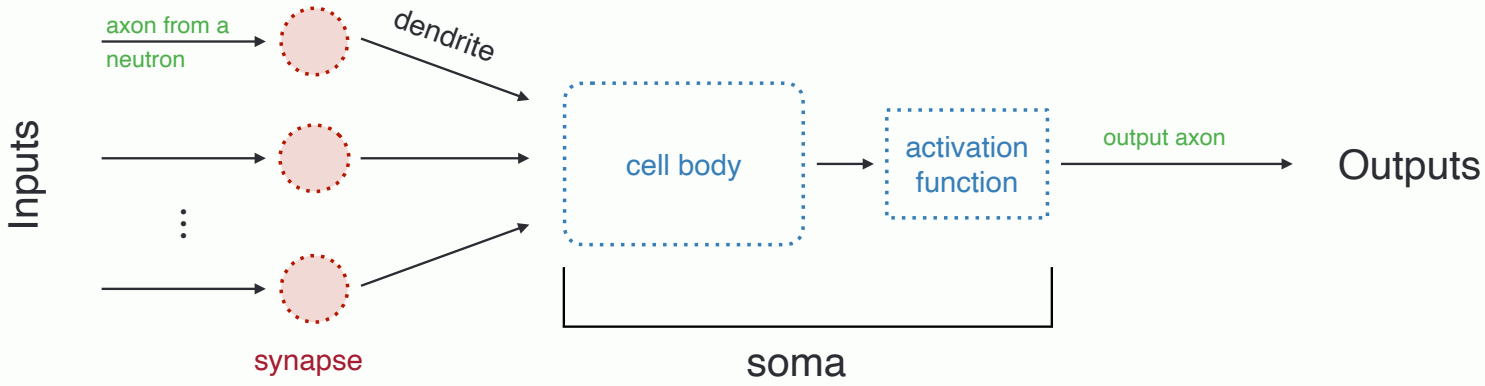
https://www.cs.toronto.edu/~lczhang/aps360_20191/lec/w02/terms.html

# ARTIFICIAL NEURON

An artificial neuron is a mathematical function conceived as a model of biological neurons.

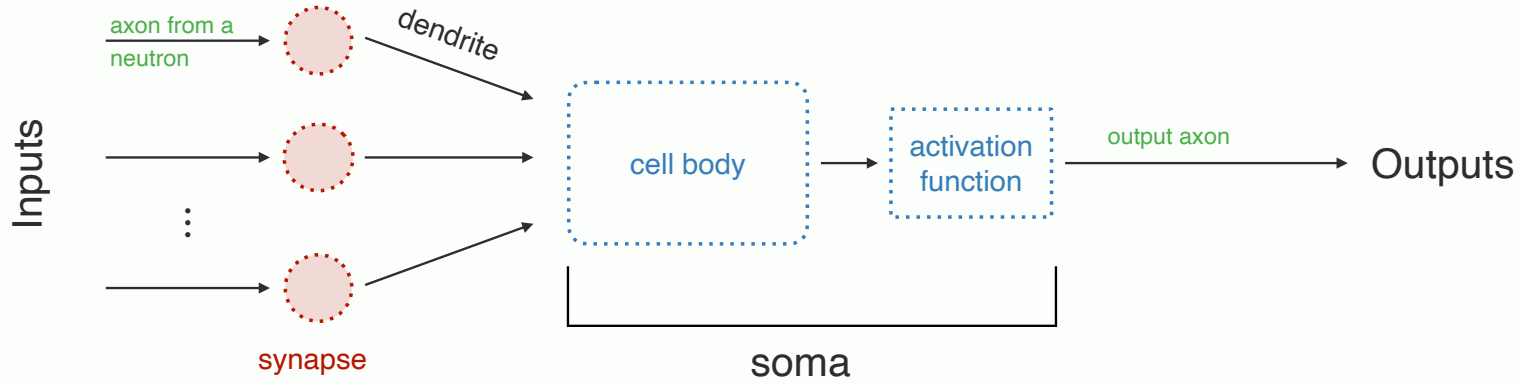
We use a simplified model of the neuron that only models the information flow.

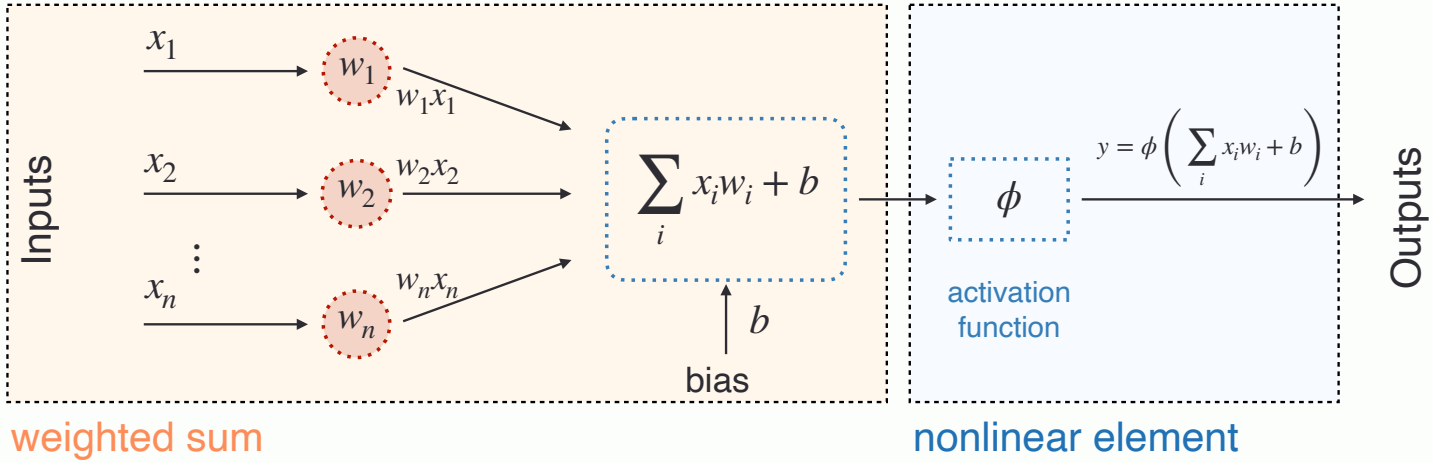



https://arxiv.org/abs/2202.01144

# ARTIFICIAL NEURON (2)



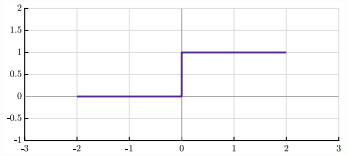
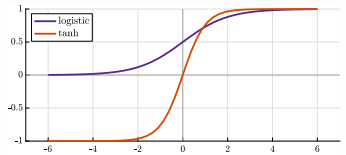
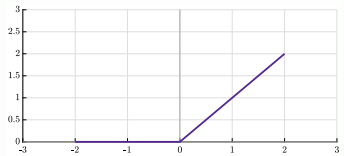

Matrix notation:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$W = [w_1, w_2, \ldots, w_n]$$

$$y = \phi\left(WX + b\right)$$

# TYPICAL ACTIVATION FUNCTIONS

| | Function $g(x)$ | Derivative | |
|---|---|---|---|
| Binary | $\begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0 \end{cases}$ | $0$ |  |
| Identity | $x$ | $1$ |  |
| Logistic/sigmoid | $\frac{1}{1+e^{-x}}$ | $g(x)(1-g(x))$ |   |
| Hyperbolic tangent | $\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$ | $1-g^2(x)$ | |
| ReLU | $\begin{cases} 0 & \text{if } x \leq 0, \\ x & \text{if } x > 0 \end{cases}$ | $\begin{cases} 0 & \text{if } x < 0, \\ \text{undefined} & \text{if } x = 0, \\ 1 & \text{if } x > 0 \end{cases}$ |  |
| Gaussian | $e^{-x^2}$ | $-2xe^{-x^2}$ |  |

https://en.wikipedia.org/wiki/Activation_function

# From a single neuron to ANNs

# DEFINITION OF ANN

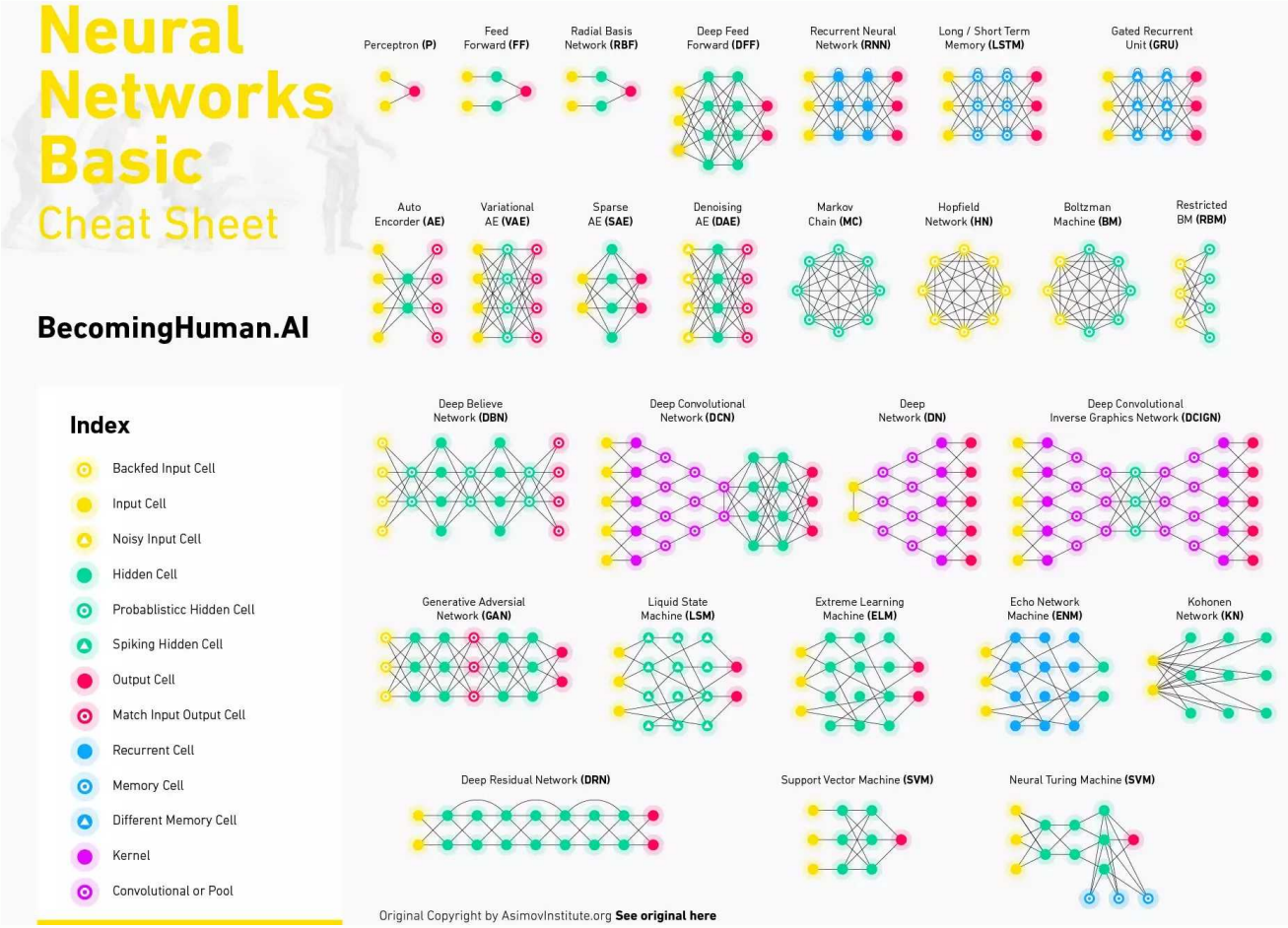
… *a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs*.

Dr. Robert Hecht-Nielson as quoted in “Neural Network Primer: Part I” by Maureen Caudill, AI Expert, Feb. 1989

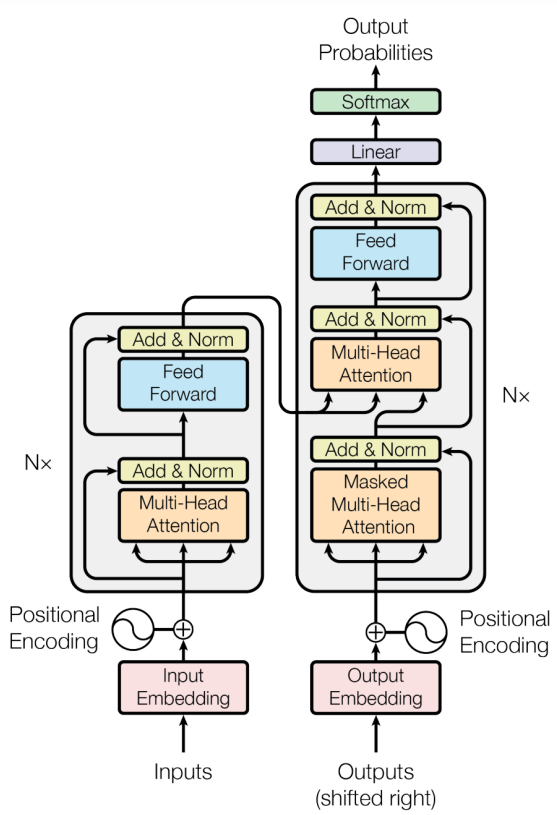
# HISTORY

- 1943 - Warren McCulloch and Walter Pitts introduced the first model of a neural network, which was based on the biological neurons in the brain.
- 1950s - Frank Rosenblatt developed the Perceptron, a neural network that could learn to recognize patterns in data.
- 1960s and 1970s - neural networks became less popular due to the limitations of the hardware available at the time.
- 1980s - backpropagation algorithm was developed, which made it possible to train neural networks with more than one layer. This led to a resurgence of interest in neural networks.
- Early 2000s - deep learning emerged as a new approach to building neural networks with multiple layers
- Nowadays …

# SOME STRUCTURES …

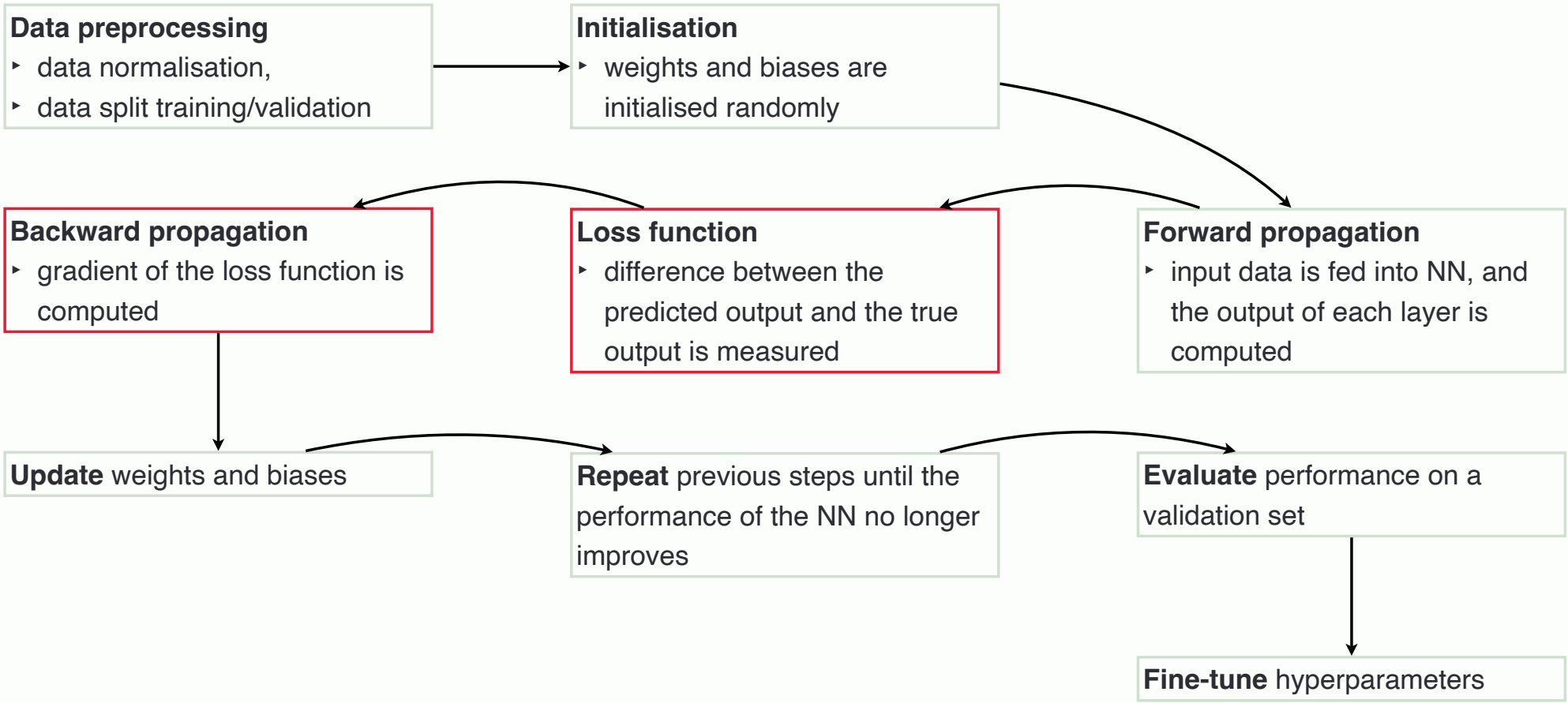



## Transformers




https://becominghuman.ai/cheat-sheets-for-ai-neural-networks-machine-learning-deep-learning-big-data-science-pdf-f22dc900d2d7
https://arxiv.org/abs/1706.03762v5

# TRAINING: GENERAL FLOW



# BIG RACE IN AI

# TYPICAL APPLICATION DOMAINS

| | | |
|---|---|---|
| GAMES | CREATIVE | HOME & LIFESTYLE |
| PROFESSIONAL | FINANCE | ADMINISTRATIVE |
| INDUSTRIAL | AGRICULTURAL | TRANSPORT |
| GOVERNMENT | EDUCATION | MEDICINE |
| SCIENCE | SECURITY | COMPUTING |
| SPEECH & LANGUAGE | VISION | ROBOTICS |

# CLUSTERING

## What is it?

Unsupervised machine learning technique that groups unlabeled examples based on their similarity to each other.

Clustering
- uses **unlabeled data** and looks for similarities between groups (clusters) and
- attempt to segment the data into separate clusters.

Note!
- Clustering doesn't provide any labels to the groups
- If groups are labeled, it is called classification

# BASIC IDEA OF CLUSTERING

## Clustering 2-dimensional data

| X1 | X2 |
|---|---|
| 2 | 4 |
| 2 | 2.5 |
| 3 | 3 |
| 3 | 5 |
| 4 | 2 |
| 4 | 4 |
| 6 | 7 |
| 7 | 6 |
| 8 | 5 |
| 8 | 7 |
| 9 | 6 |
| 9 | 7 |

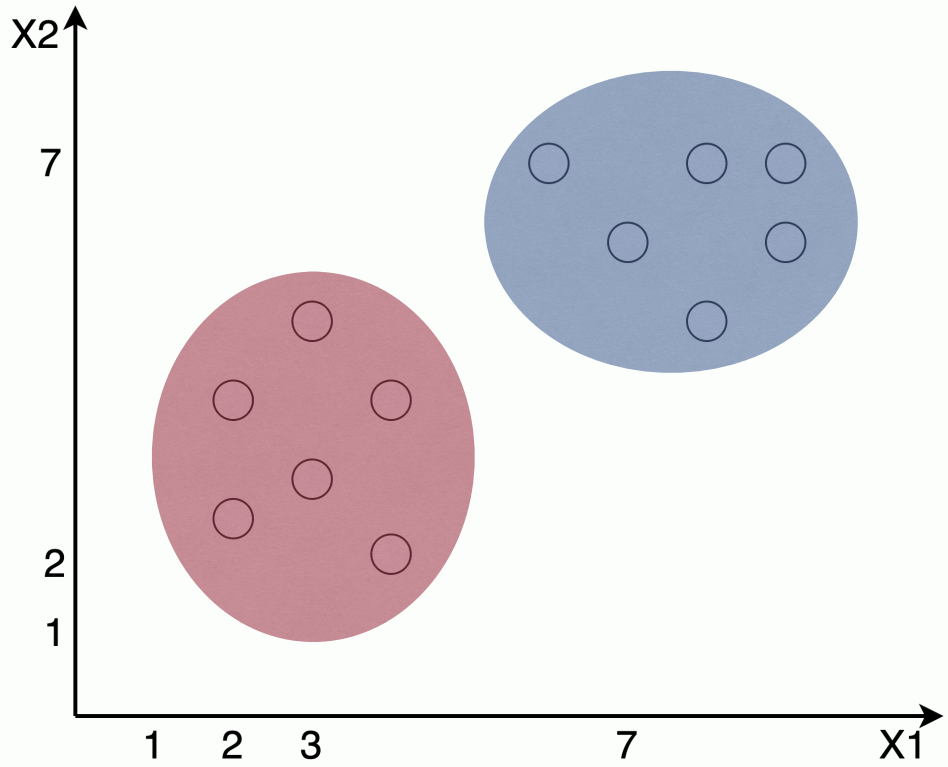
- We have only features (input data)
- How could we cluster this data together?
- Let's plot the data

# BASIC IDEA OF CLUSTERING (2)

## Clustering 2-dimensional data

| X1 | X2 |
|---|---|
| 2 | 4 |
| 2 | 2.5 |
| 3 | 3 |
| 3 | 5 |
| 4 | 2 |
| 4 | 4 |
| 6 | 7 |
| 7 | 6 |
| 8 | 5 |
| 8 | 7 |
| 9 | 6 |
| 9 | 7 |

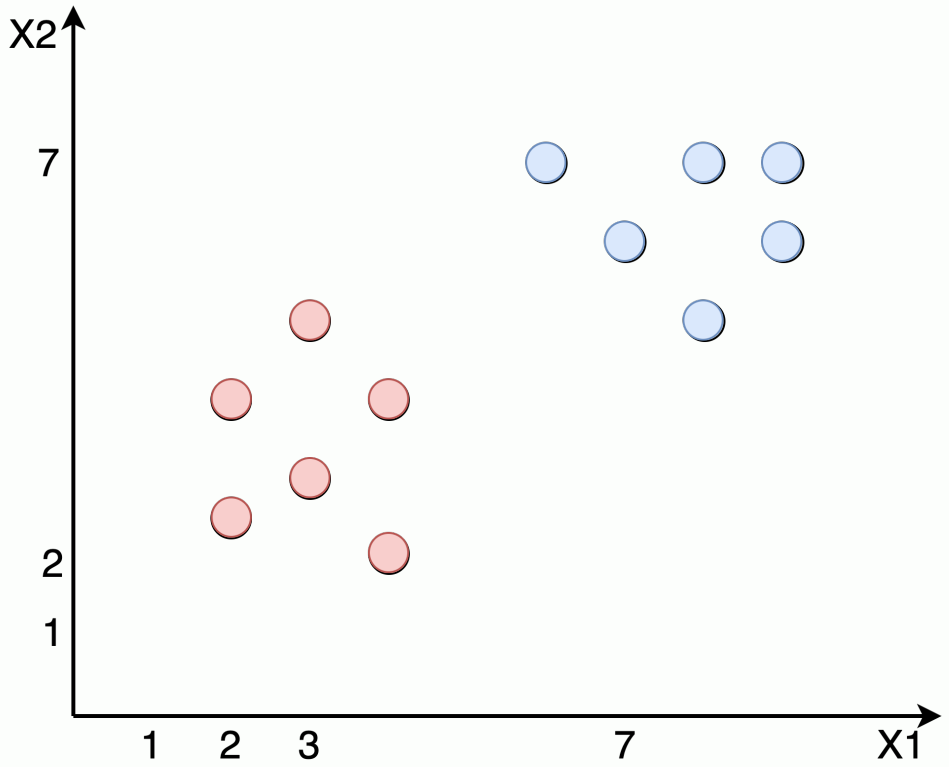
Distance is the intuitive metric!



# BASIC IDEA OF CLUSTERING (3)

## Clustering 2-dimensional data

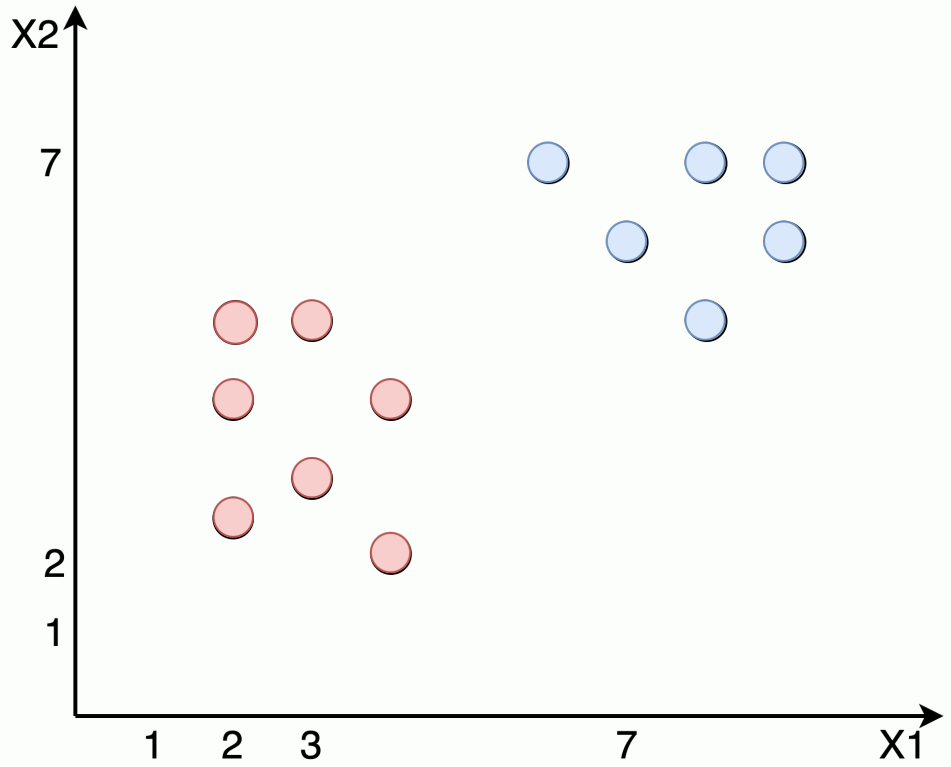
| X1 | X2 |
|---|---|
| 2 | 4 |
| 2 | 2.5 |
| 3 | 3 |
| 3 | 5 |
| 4 | 2 |
| 4 | 4 |
| 6 | 7 |
| 7 | 6 |
| 8 | 5 |
| 8 | 7 |
| 9 | 6 |
| 9 | 7 |



# BASIC IDEA OF CLUSTERING (4)

## Clustering a new pattern

| X1 | X2 |
|---|---|
| 2 | 5 |

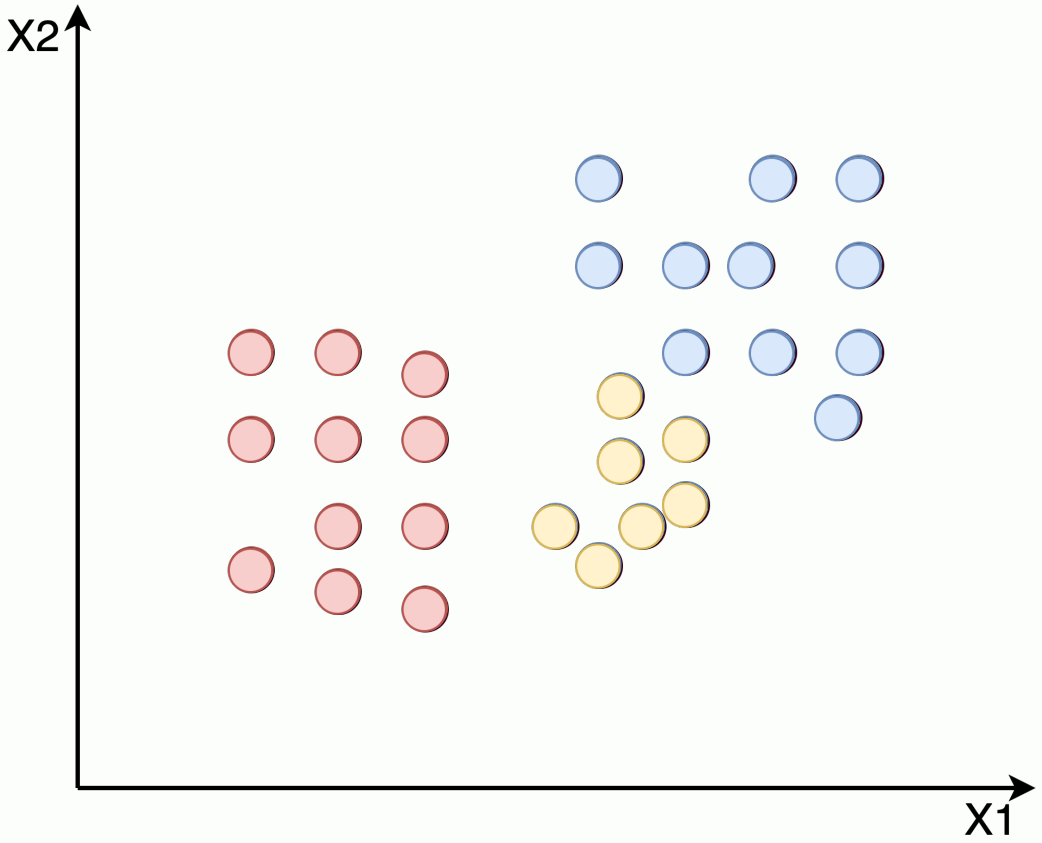


# MAIN CHALLENGES

- Actually, we don’t know for sure if this was a correct way of grouping together these data points or not.
- There were no correct labels to validate the result!
- What about situations that are not so obvious or multi-dimensional?

# MAIN CHALLENGES

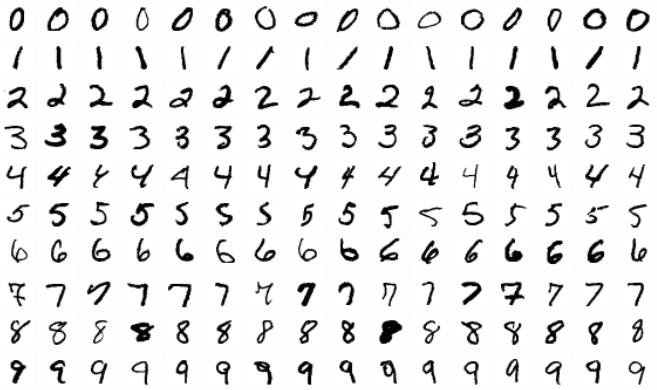
## How many clusters?



# CLUSTERING OF PATTERNS

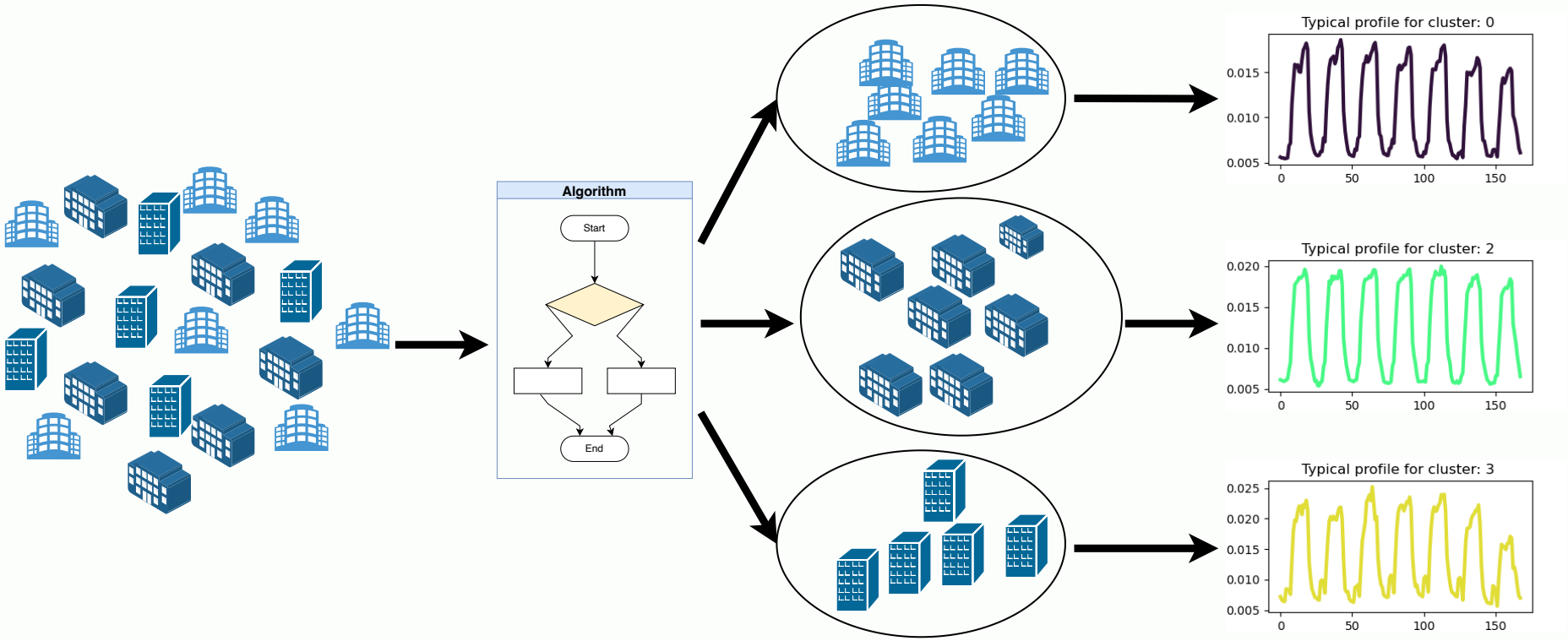
## On example of small images (handwritten digits)

**MNIST** database of handwritten digits
- 26x26px
- 60 000 labeled images and 10 000 test images



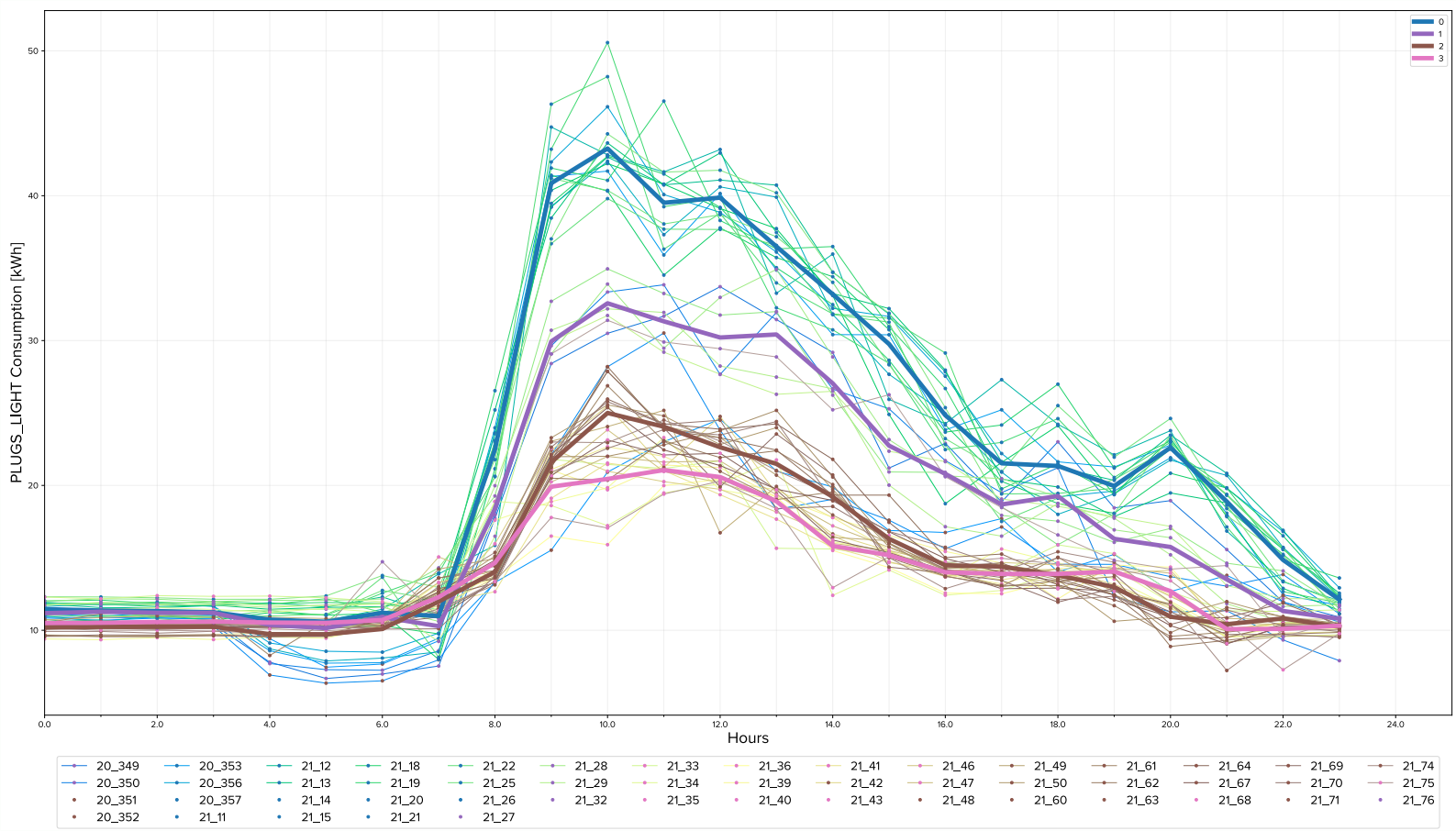
**Semeion**
- 1593 handwritten digits from around 80 persons
- 16x16px

# CLUSTERING BUILDINGS BY ENERGY PERFORMANCE



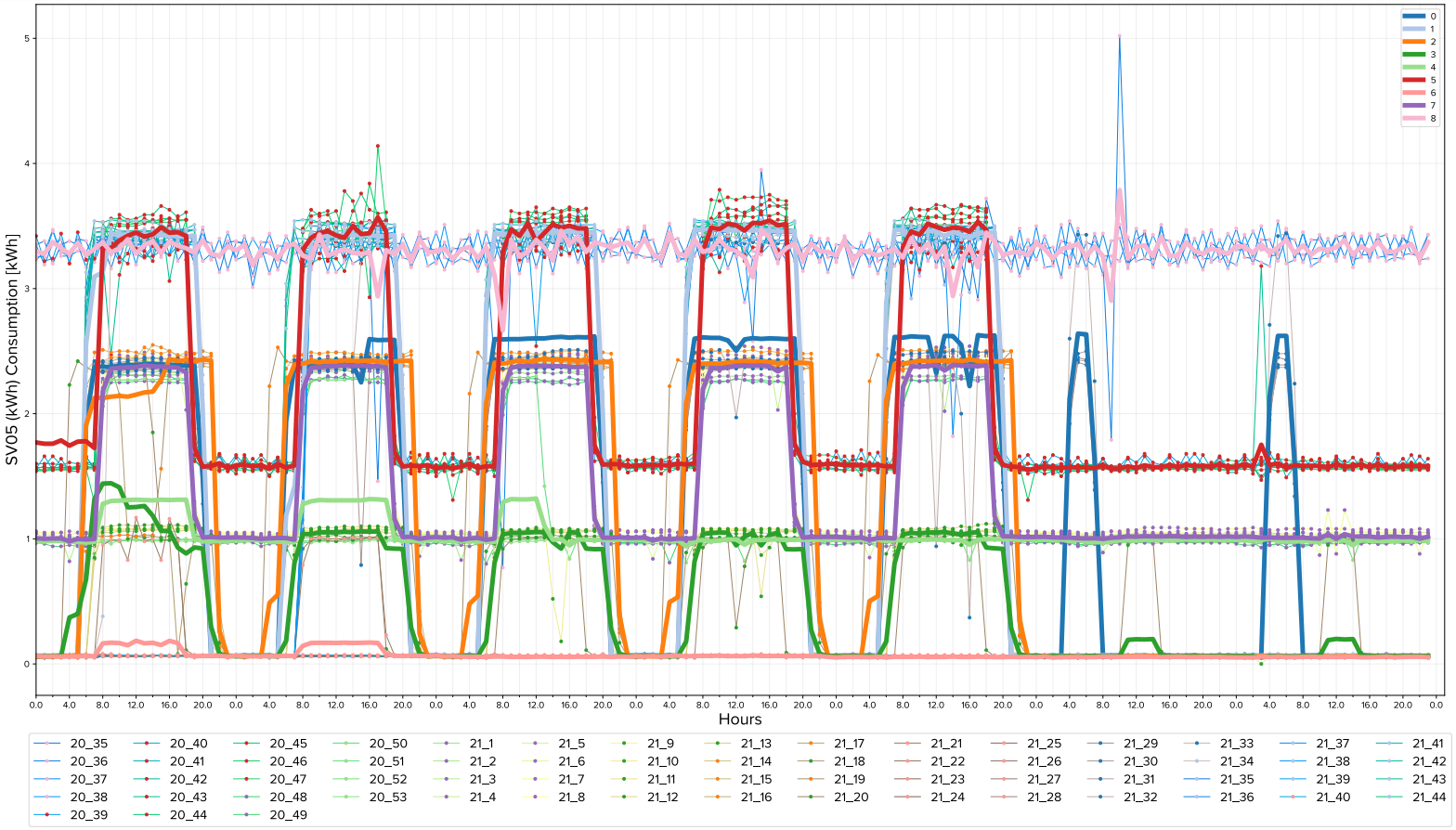
# CLUSTERING BUILDINGS BY ENERGY PERFORMANCE

## Daily consumption of plugs and light in a building



# CLUSTERING BUILDINGS BY ENERGY PERFORMANCE

## Weekly consumption of a ventilation unit in a building



# DIGITAL TRANSFORMATION CHALLENGES

| | |
|---|---|
| Data availability | Collection of data from different sources and machines requires intensive number of sensors and inter-connectivity. |
| Data quality and accessibility | Failure to access high quality data will mislead the systems and models. |
| Data transmission | High speed and reliable communication are necessary for transferring vital data |
| Trojan and poisoning attacks | Manipulation of data may lead to significant loss in making decisions. |
| Securing communications | Unsecure communications can have unpredicted results for system. |
| Trustworthiness | Humans need to trust the machine produced results and predictions. |
| Computational cost | Most algorithms are complicated for deployment in real applications. |
| Real time performance | Some critical sections demand for real time decision makings and interpretation. Like autonomous drivers. |
| Black box AI | Majority of AI models are not clearly explainable |
| Unmature technologies | Technologies like block chain are still unmature for real world applications |
| User privacy | Collecting data specially related to society and people always has challenges related to user privacy |

# BENEFITS OF AI-DRIVEN CONTROL

- Analyzes huge amounts of data (“Big Data”)
- Can model and predict complex dynamics
- Can make forecasts and be proactive
- Considers many different parameters
- Can learn
- Significant improvement in performance without additional hardware
- Can synchronizes many sub-components

**As the result…**

- Improved performance

# CHALLENGES RELATED AI-DRIVEN CONTROL

- Explainability: Black box and grey box models
- Complexity
- Computational costs
- Data quality, availability and amount
- Trustworthiness
- Safety
- Reliability
- Stability

# Thank you!

Energy Efficiency Centre of Excellence
ENERGIATÕHUSUSE TIPPKESKUS

TAL TECH