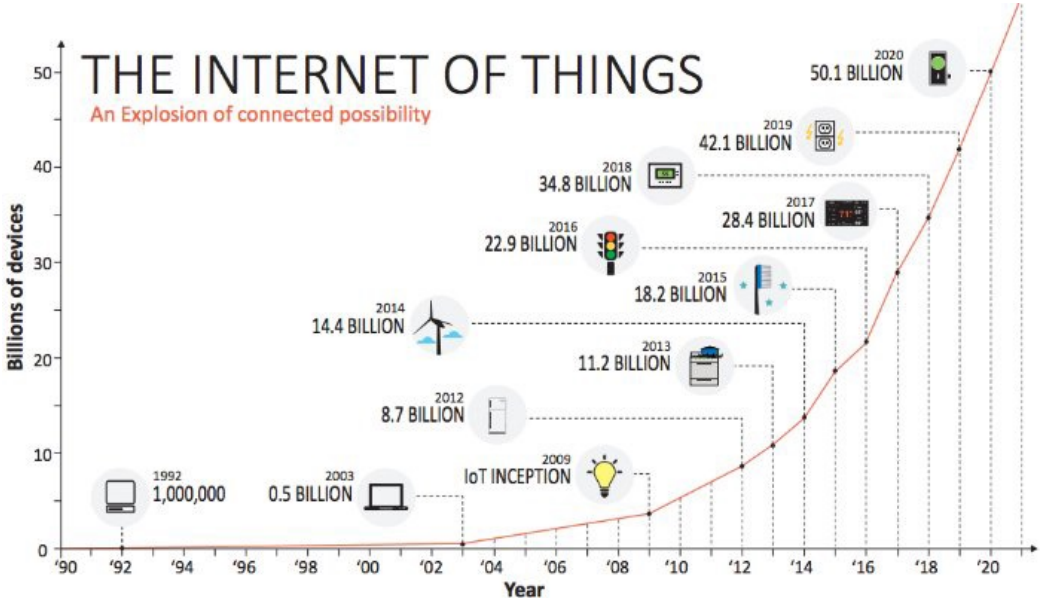
Edge Computing

Master 's Degree in Informatics Engineering

Current situation – Internet of Things

The number of interconnected devices is growing exponentially



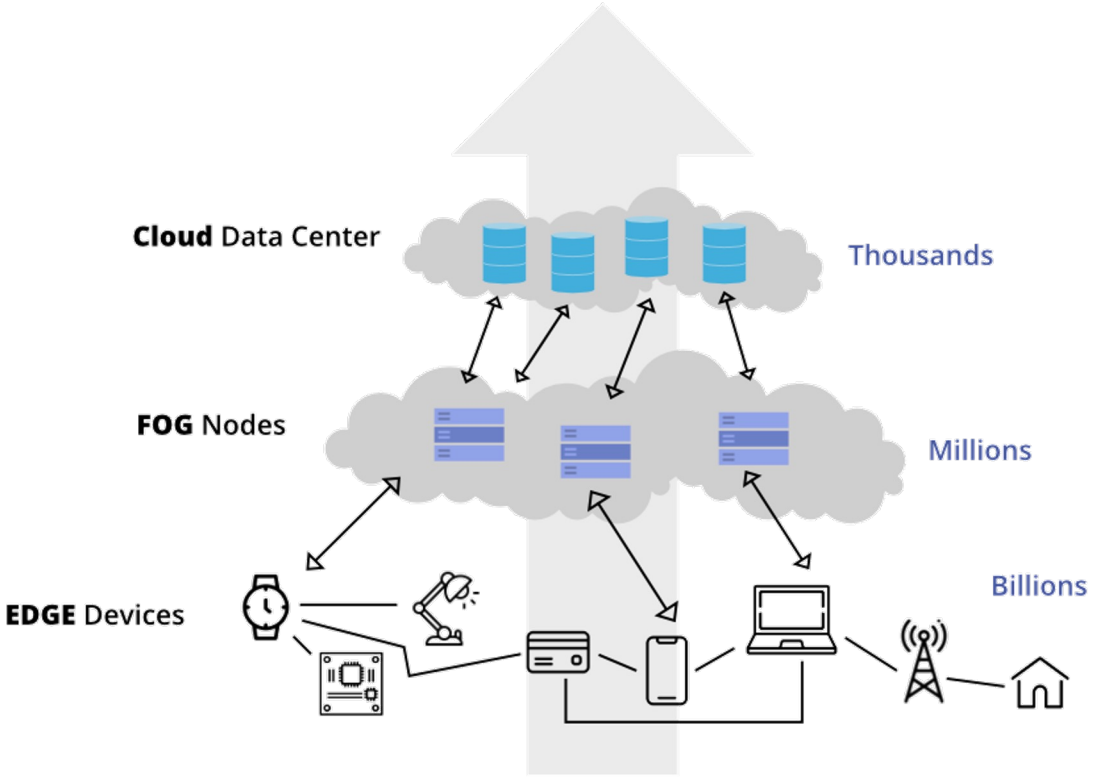
- From 9.7 billion in 2020 to more than 29 billion IoT devices in 2030.
- The total amount of data created, captured, copied, and consumed globally is forecast to increase rapidly, reaching 64.2 zettabytes (10^{21}) in 2020. In 2025 is projected to grow to more than **180** zettabytes.
- This is 277 times greater than the amount of data end-user devices will transmit to data centers, and 47 times greater than projected total data center traffic.

https://www.statista.com/

Current situation – Internet of Things

Open questions:

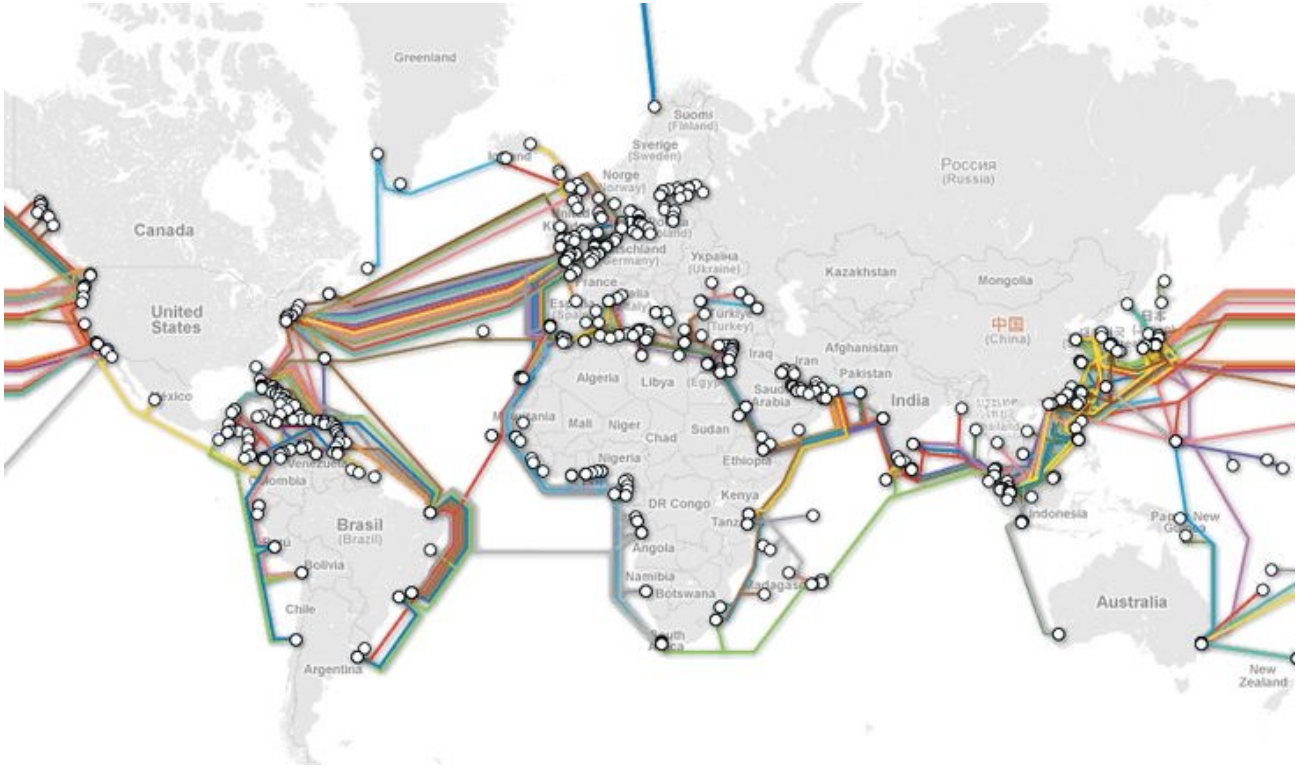
1. How to perform effective analysis on this huge amount of data.
2. All data is really necessary?



Current situation – Internet of Things

Open questions:

1. How to perform effective analysis on this huge amount of data.
2. All data is really necessary?



Current situation – Internet of Things

Open questions:

1. How to perform effective analysis on this huge amount of data.
2. All data is really necessary?

CLOUD COMPUTING SERVICE PROVIDERS

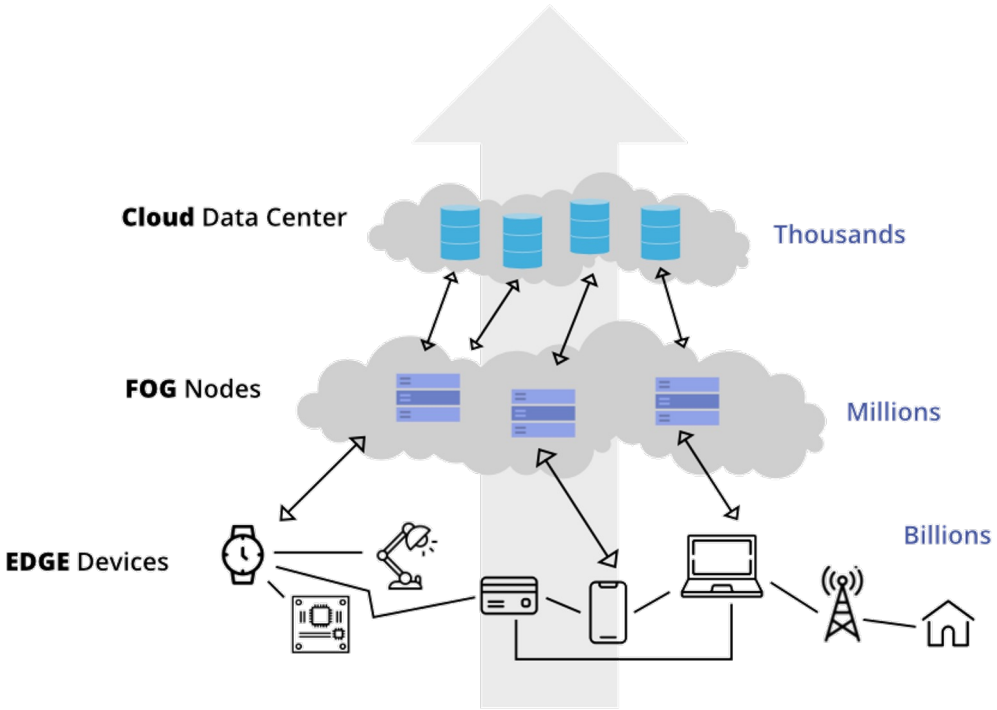
- MICROSOFT AZURE
- AMAZON WEB SERVICES
- GOOGLE CLOUD PLATFORM

Google Data Center Location

Current situation – Internet of Things

Open questions:

1. How to perform effective analysis on this huge amount of data.
2. All data is really necessary?
3. Where the analysis results are required? Data privacy



Current situation – Internet of Things

Open questions:

1. How to perform effective analysis on this huge amount of data.
2. All data is really necessary?
3. Where the analysis results are required?
4. Minimum allowable time to obtain the results? Data Latency

"To achieve a round-trip time (RTT) of less than 1 millisecond, the maximum distance between a mobile device and a cloud *can be no greater than around 100km."*

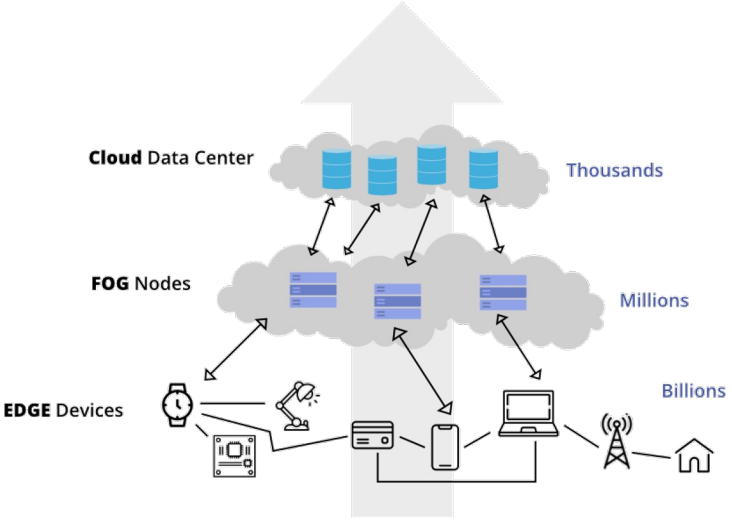
Route	Distance	Time, light in vacuum	Time, light in fiber	Round-trip time (RTT) in fiber
New York to San Francisco	4,148 km	14 ms	**21 ms**	42 ms
New York to London	5,585 km	19 ms	**28 ms**	56 ms

Comercial fibre optic > 2/3 speed of light

About Edge Computing

1. Computing paradigm that offloads computation and storage from centralized Cloud to the network's logical extremes.

2. Benefits:
   - Speed
   - Security
   - Scalability
   - Versatility
   - Faster decisions making on IoT
   - Computation and Data analysis in disconnected environments



3. Use Cases:
   - 5G Applications
   - Smart Cities
   - Self Driving Cars

Edge Computing - Devices

Small devices implies:

1. Classic situations:
    - Soft/easy devices control
    - Classification
    - Identification
    - Malfunctioning situations
2. Extremally low amount of memory

    Programs and memory usage must be optimized
3. Take control on any environment situation requires complex heuristics

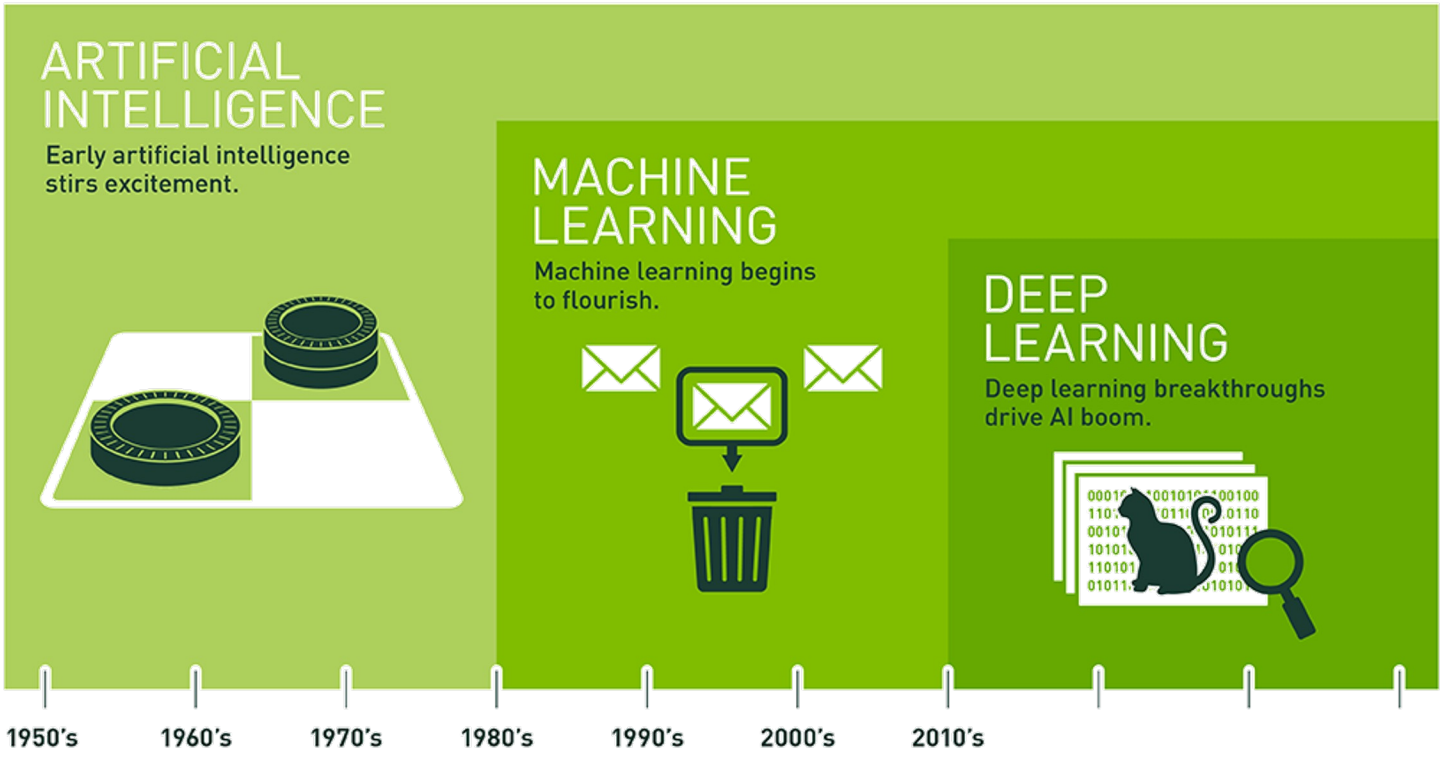
Solution

>> Machine Learning > Deep Learning

Simple algorithm with affordable computational requirements.

Deep Learning concepts

- Deep learning is a subfield of machine learning focusing on learning data representations as successive layers of increasingly meaningful representations.



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

Image from https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

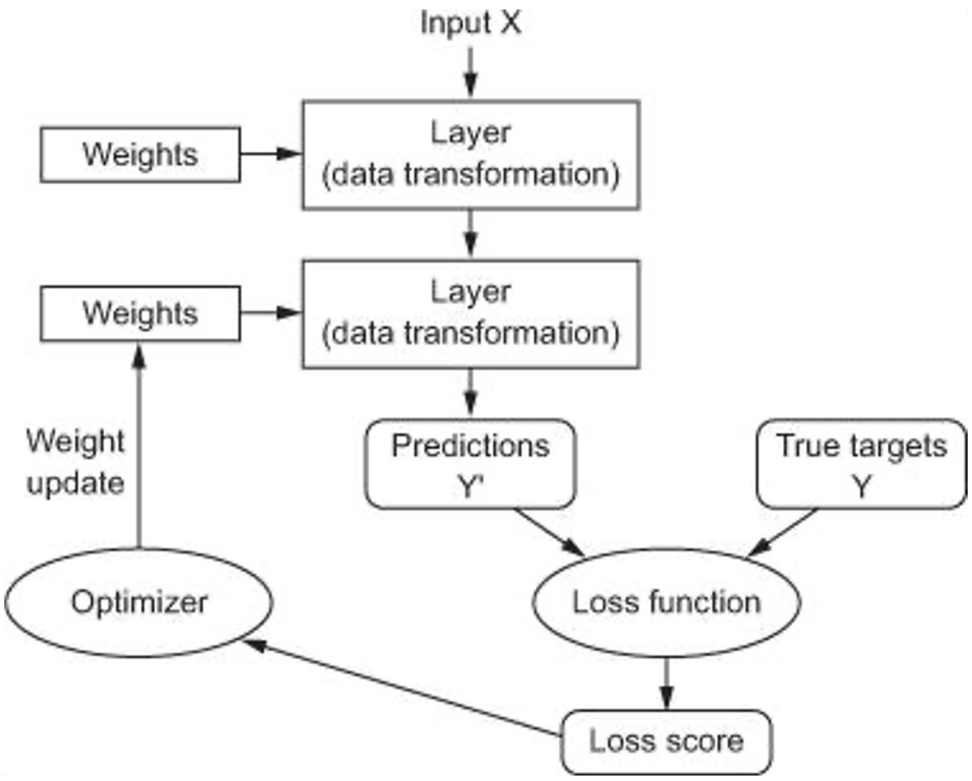
Deep Learning concepts

Layers:
- Data Processing modules
- Input/output. Multiple tensors
- Defined by states/weights combined by a network

Training process:
- Predictions are compared with the input targets

Loss function:
- Represents the acceptable accuracy

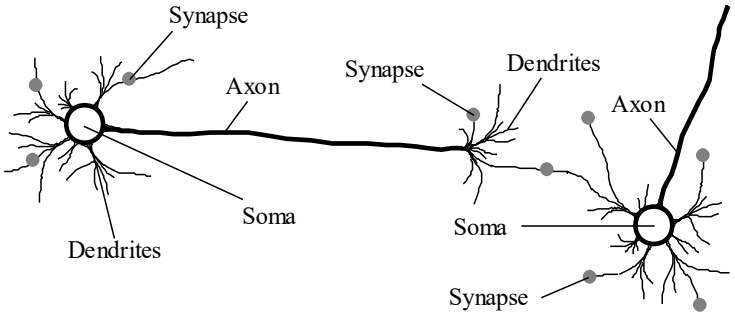
Main Process

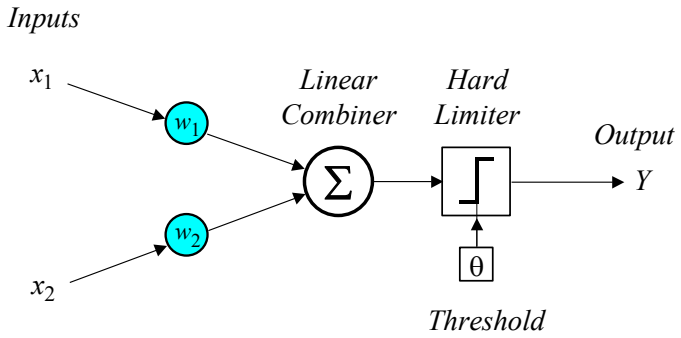


Deep Learning concepts – Neural Networks

Human neural network:

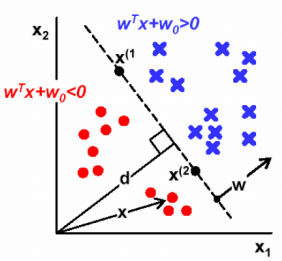
- Neuron switching time: ~ 0.001 second
- Number of neurons: ~ 10^{10}
- Connections per neuron: ~ $10^{4\sim5}$
- Scene recognition time: ~ 0.1 second



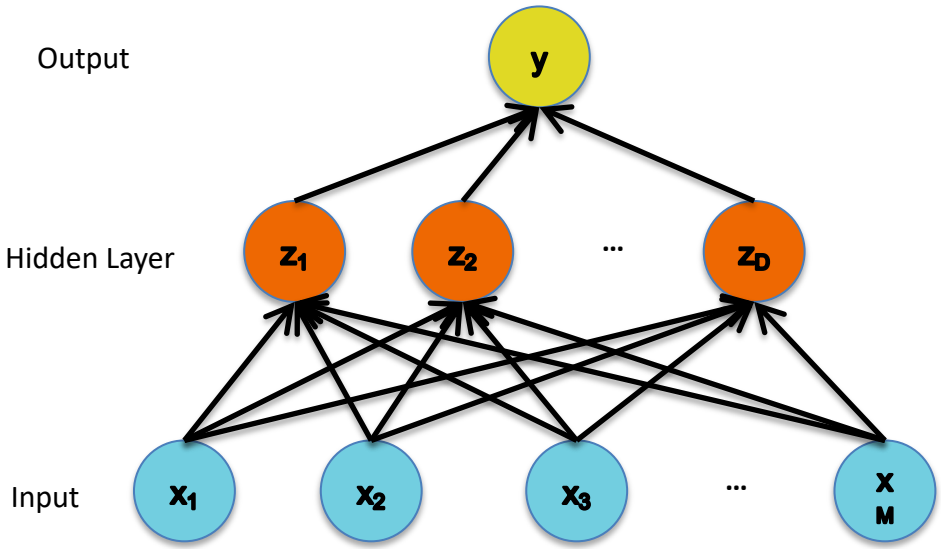


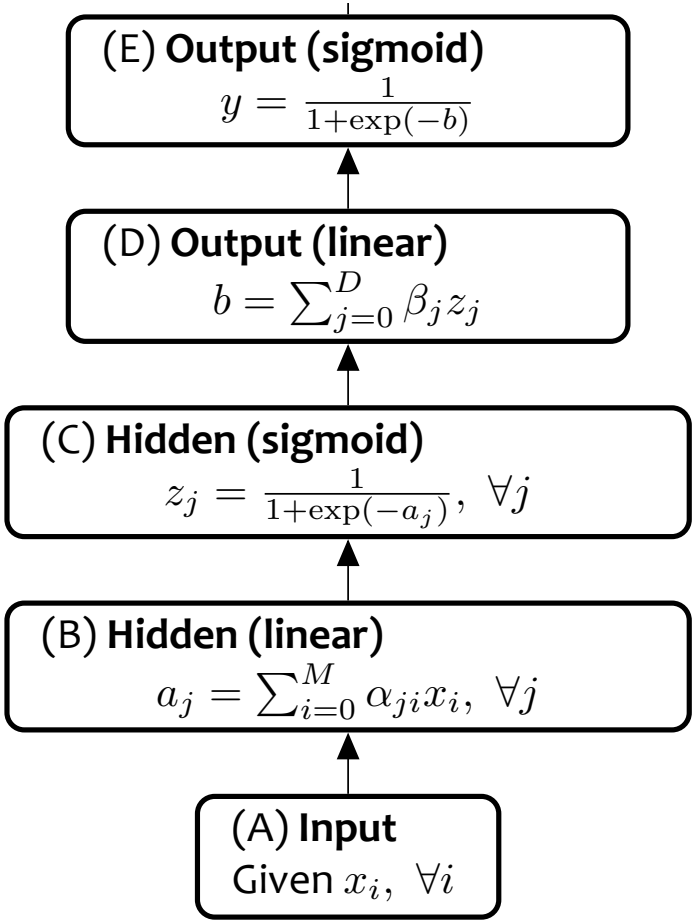
$$X = \sum_{i=1}^{n} x_i w_i$$

$$y = \begin{cases} +1, & \text{if } X \geq \omega_0 \\ -1, & \text{if } X < \omega_0 \end{cases}$$



Deep Learning concepts – Neural Networks

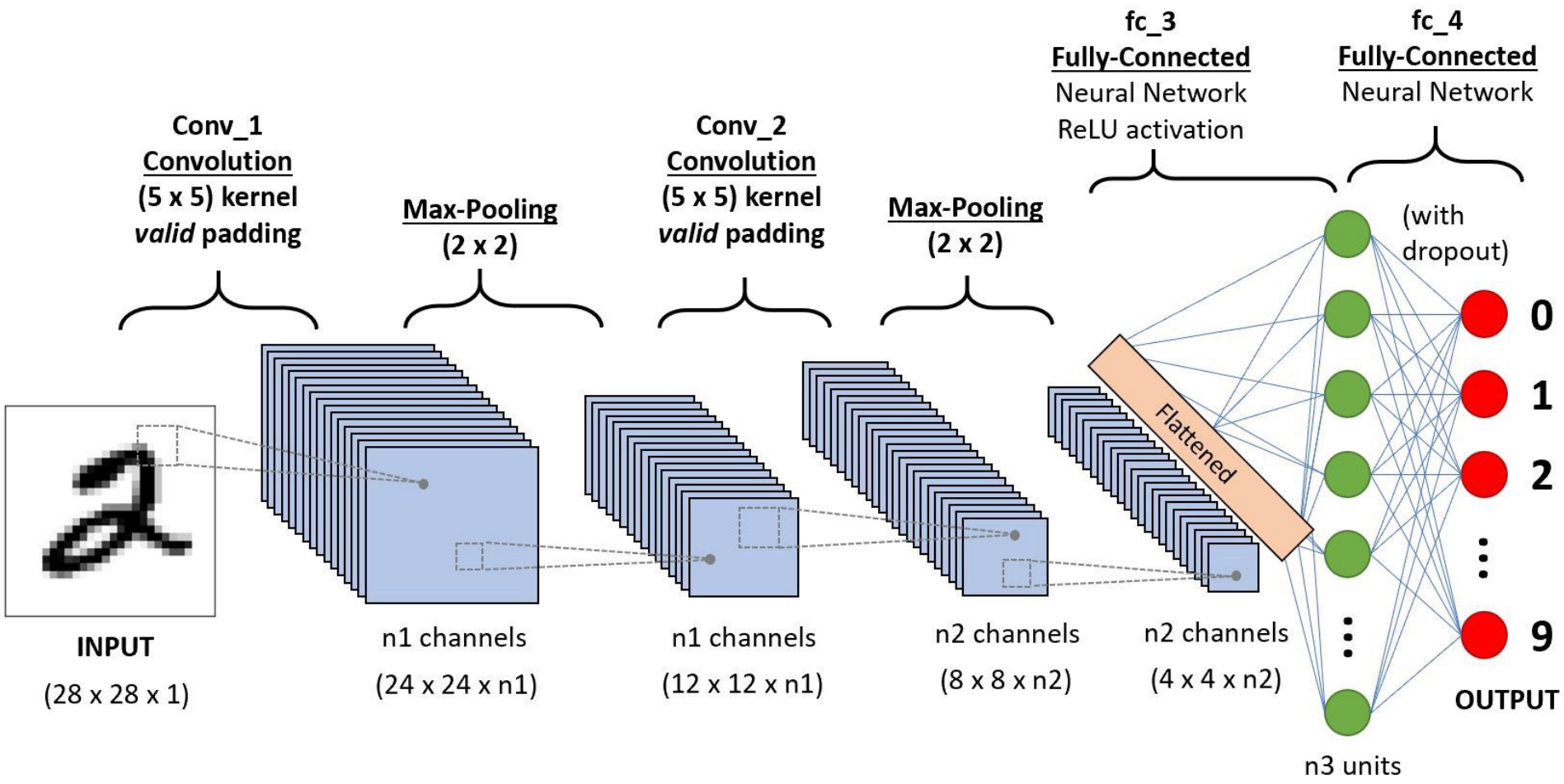




Complex architecture

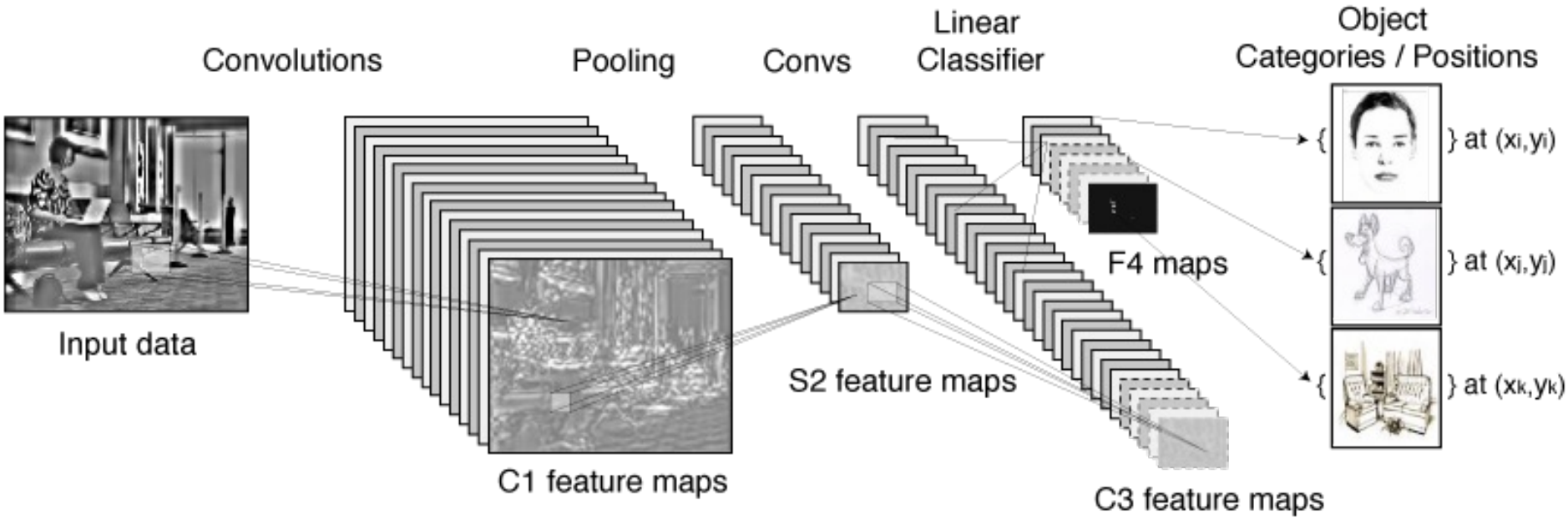
- # of hidden layers? (depth)
- # of units per hidden layer? (width)
- Type of activation function (nonlinearity)
- Form of objective function?

Deep Learning concepts – Neural Networks



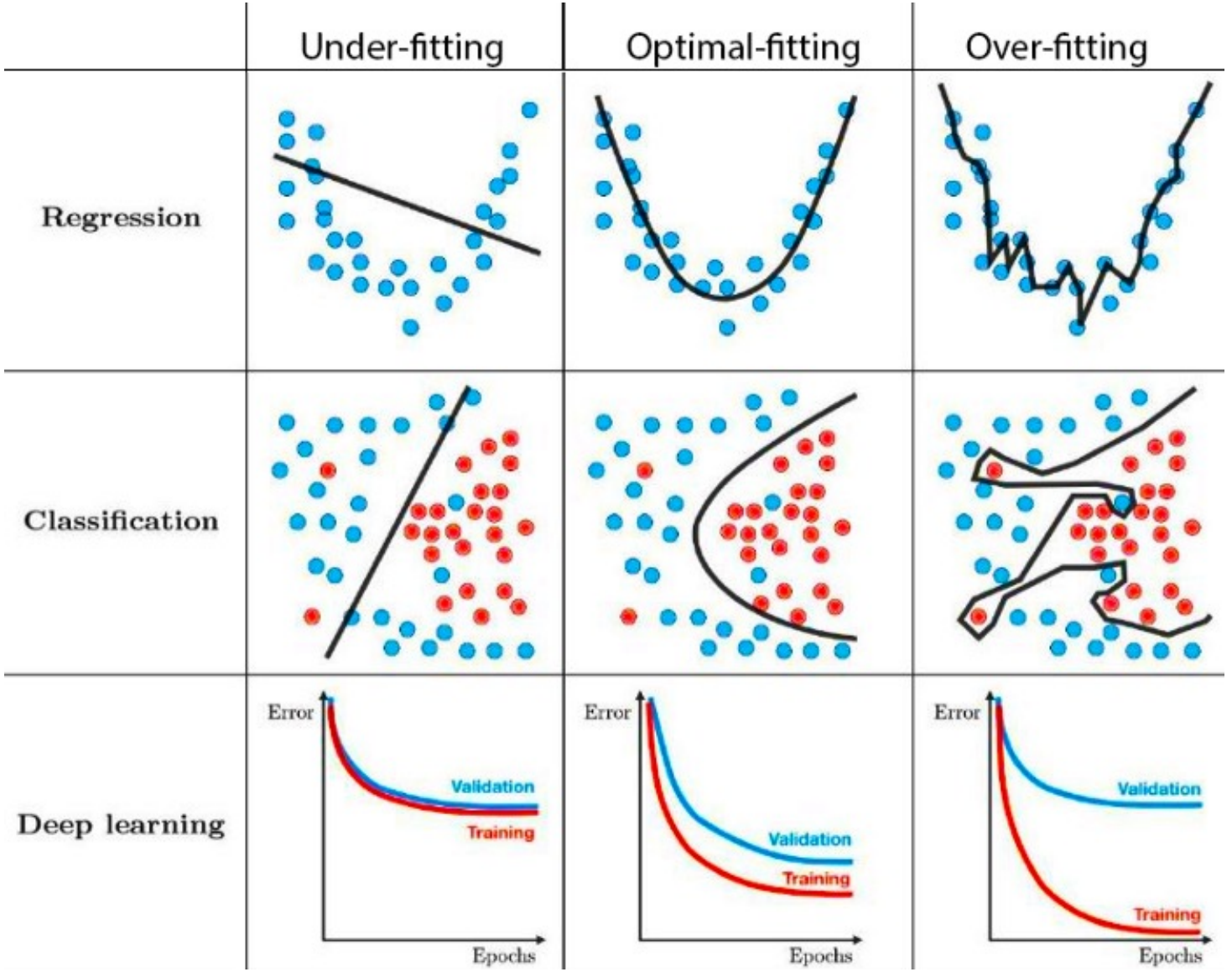
Identification/classification

Deep Learning concepts – Neural Networks



Objects recognition

Deep Learning concepts – Neural Networks



Neural Networks Attributes

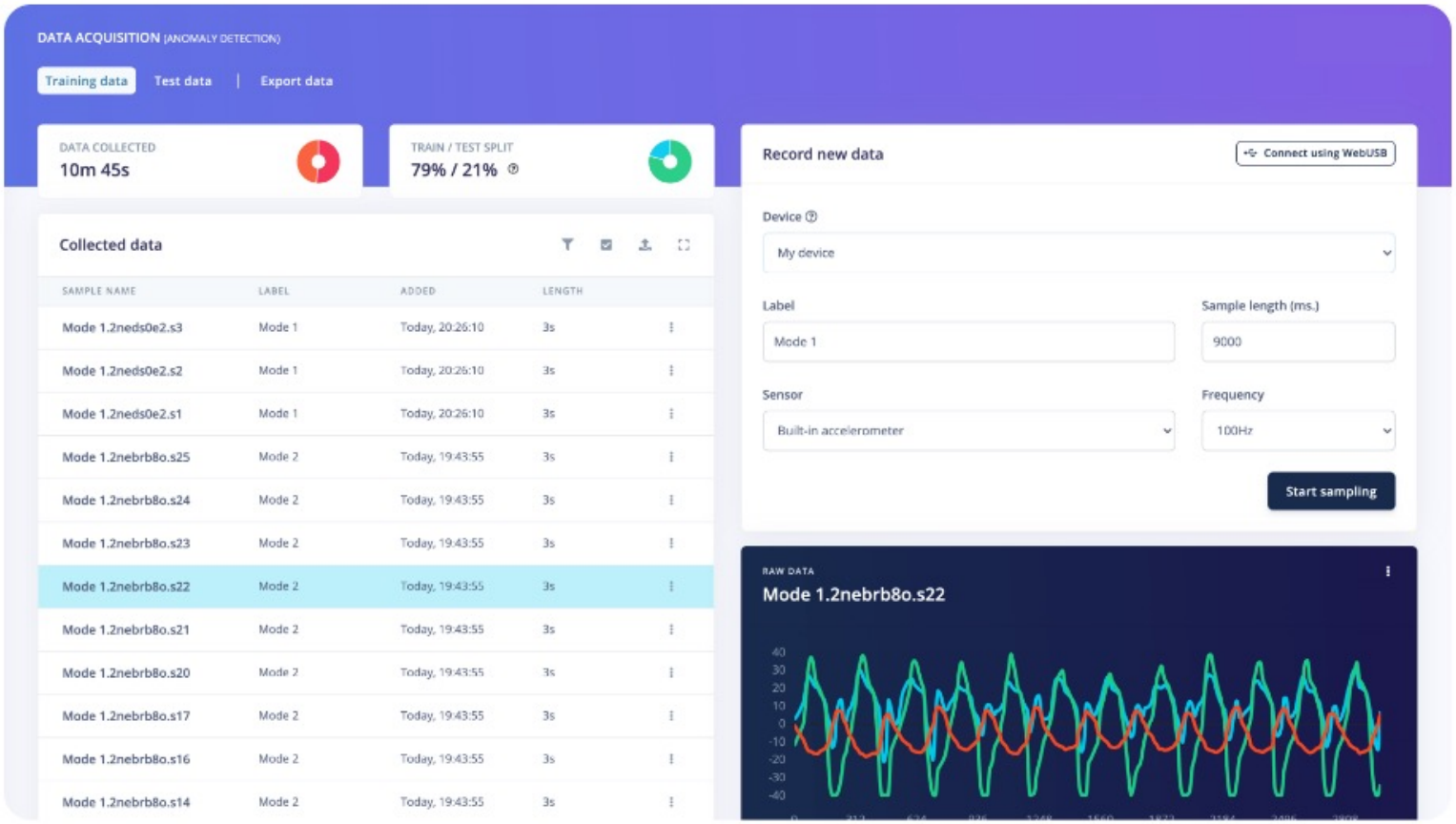
- **Adaptive Learning**: Neural networks model non-linear and complex relationships and build on previous knowledge.
- **Self-Organization**: It is possible to redefine the inner structure (synapsis) allowing to cluster and classify vast amounts of data.
- **Real-Time Operation**: Neural networks can (sometimes) provide real-time answers (they are based on basic computational operations).
- **Prognosis**: Ability to predict based on models.
- **Fault Tolerance**: When significant parts of a network are lost or missing, neural networks can redefine their organization and refill the blanks.

>> Main drawback – It requires a learning process and in some cases reinforcement learning

Neural Networks – Main use

- **Event prediction and simulations**: Production of expected output values based on incoming data.
- **Recognition and classification**: Association of patterns and organization of data sets in predefined classes. Even identifying unique features without prior data.
- **Data processing and modelling**: Data validation, aggregation and analysis. Design and troubleshooting in complex software systems.
- **Control engineering**: Monitoring of computer systems and manipulation of robots. Including the creation of autonomous systems and robots.
- **Artificial Intelligence**: Forming part of the deep learning and machine learning technologies that are fundamental parts of artificial intelligence

EDGE IMPULSE



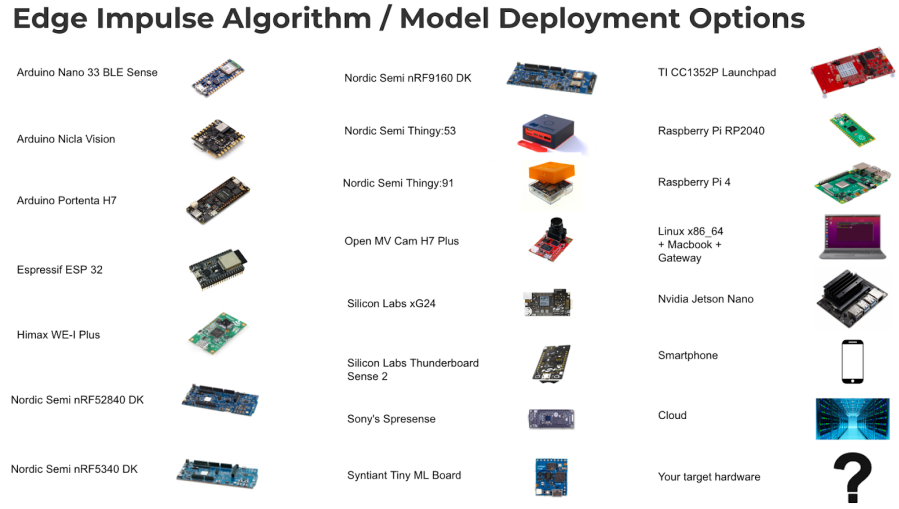
Edge Impulse

Development platform for machine learning on edge devices, free for developers and trusted by enterprises.

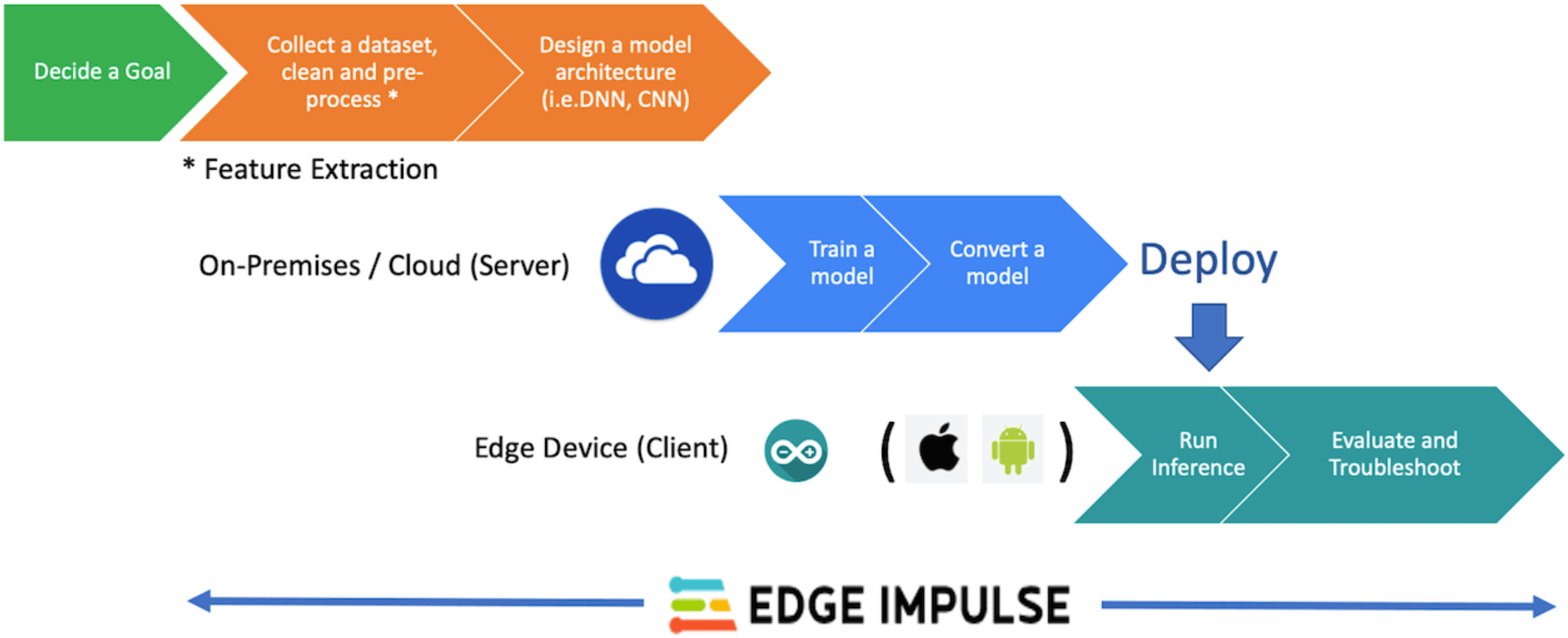
https://www.edgeimpulse.com/

https://docs.edgeimpulse.com/docs

It is based on TinyML that allows to apply machine learning techniques on embedded devices.



Edge Impulse Workflow



Edge Impulse Example

Building a Fall Detection System with Edge Impulse and a Notecard

https://blues.io/blog/fall-detection/

Japan has 36 million people aged 65 or older, and the number of fall-related deaths has increased in recent years. To help address this, Naveen Kumar built a device that detect falls with accelerometer data, and sends emergency notifications using a cellular Notecard.