WALC 2023
Applied AI

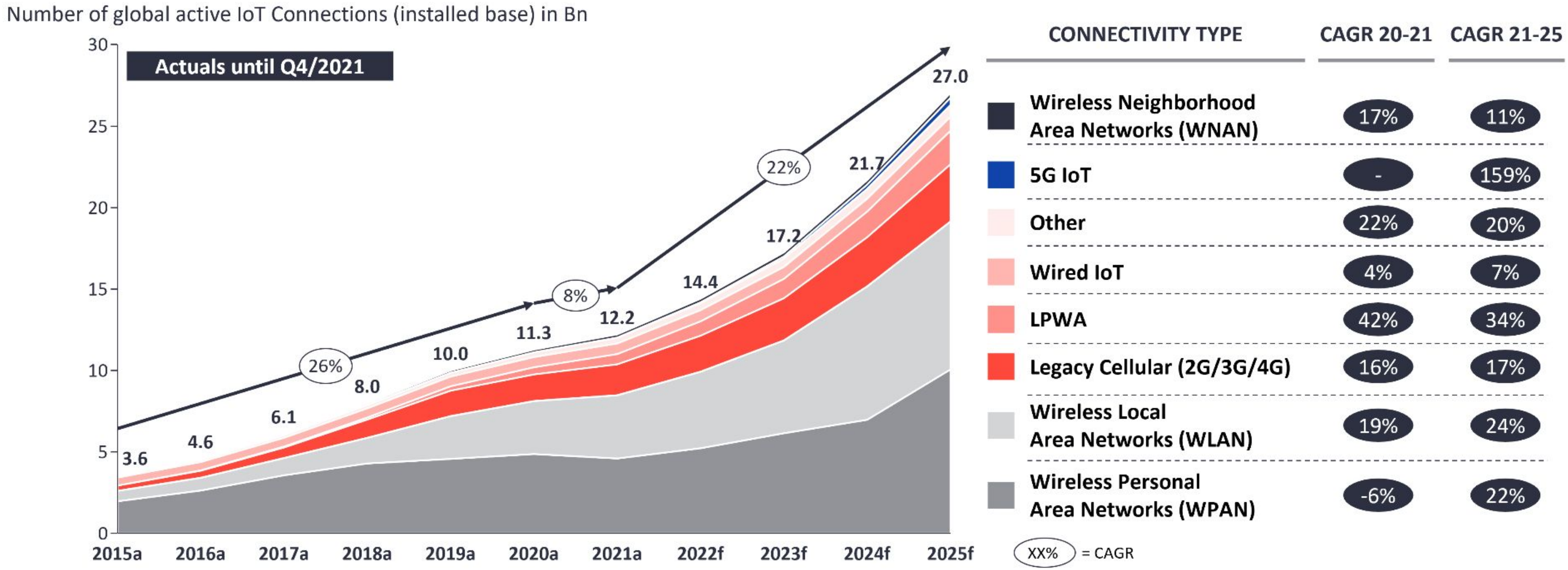
Embedded ML (TinyML) Intro & Applications

Prof. Marcelo J. Rovai
rovai@unifei.edu.br

UNIFEI - Federal University of Itajuba, Brazil
TinyML4D Academic Network Co-Chair

Internet of Things (IoT)

Global IoT Market Forecast [in billion connected IoT devices]

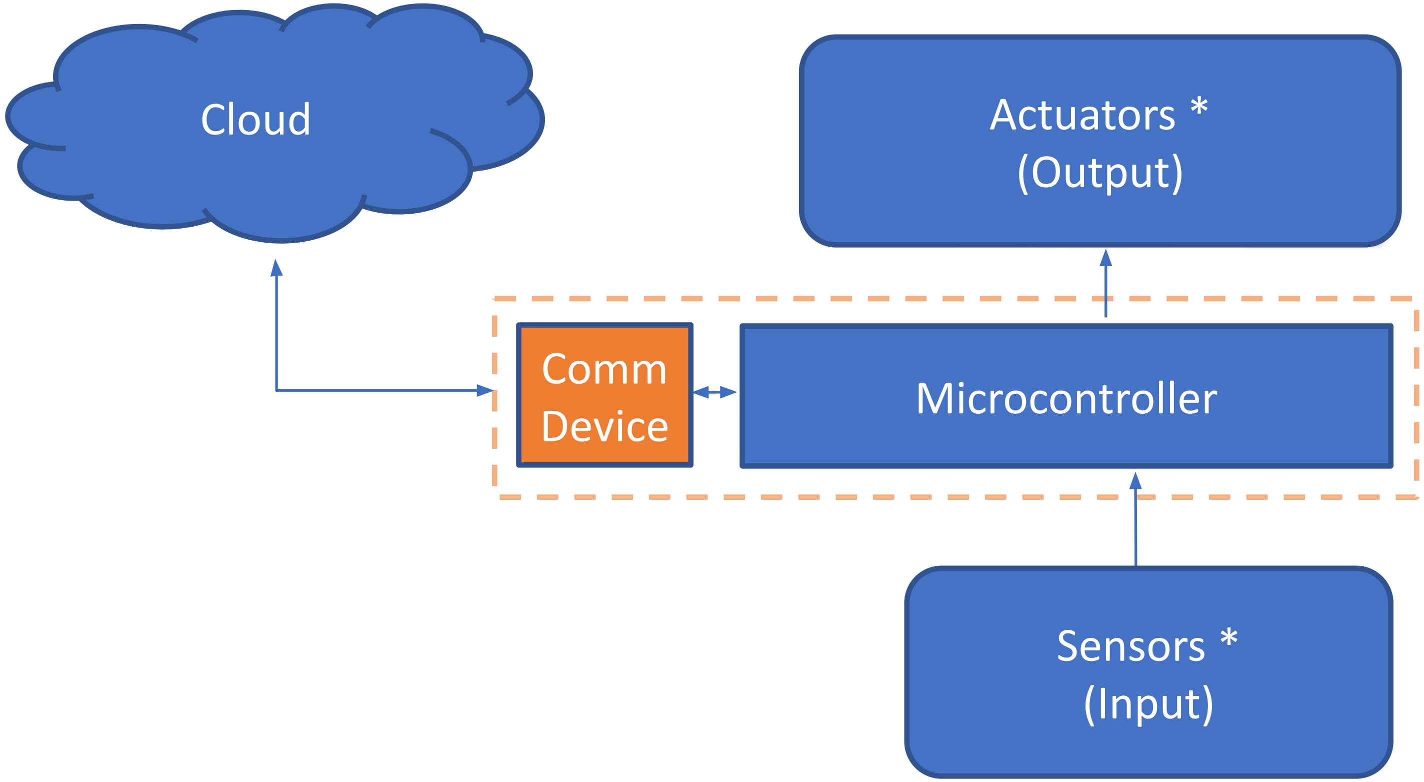


Note: IoT Connections do not include any computers, laptops, fixed phones, cellphones or tablets. Counted are active nodes/devices or gateways that concentrate the end-sensors, not every sensor/actuator. Simple one-directional communications technology not considered (e.g., RFID, NFC). Wired includes Ethernet and Fieldbuses (e.g., connected industrial PLCs or I/O modules); Cellular includes 2G, 3G, 4G; LPWAN includes unlicensed and licensed low-power networks; WPAN includes Bluetooth, Zigbee, Z-Wave or similar; WLAN includes Wi-fi and related protocols; WNAN includes non-short range mesh, such as Wi-SUN; Other includes satellite and unclassified proprietary networks with any range.

Source: IoT Analytics Research 2022. We welcome republishing of images but ask for source citation with a link to the original post and company website.

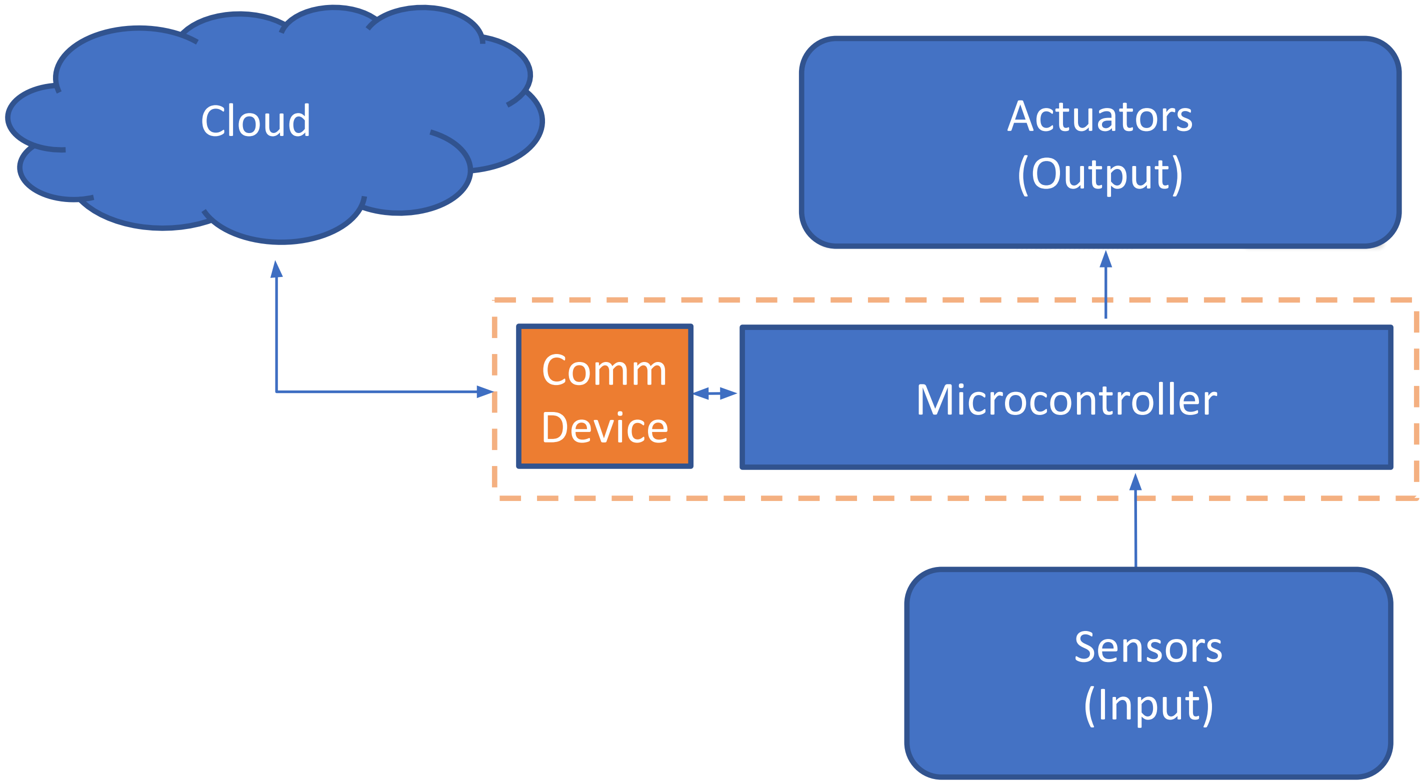
https://iot-analytics.com/number-connected-iot-devices

Typical IoT Project



* "Things"

Typical IoT Project



5 Quintillion

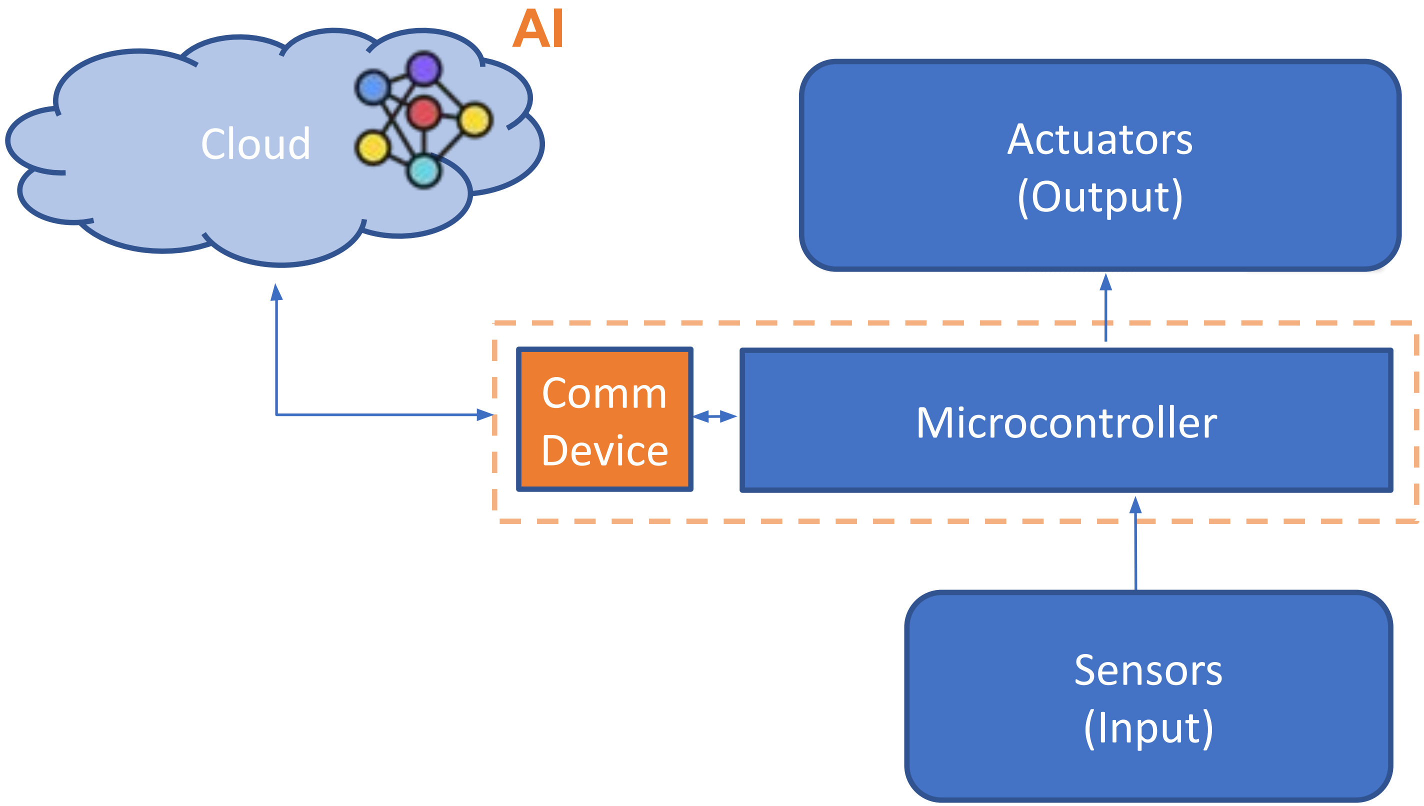
bytes of data produced every day by IoT

<1%

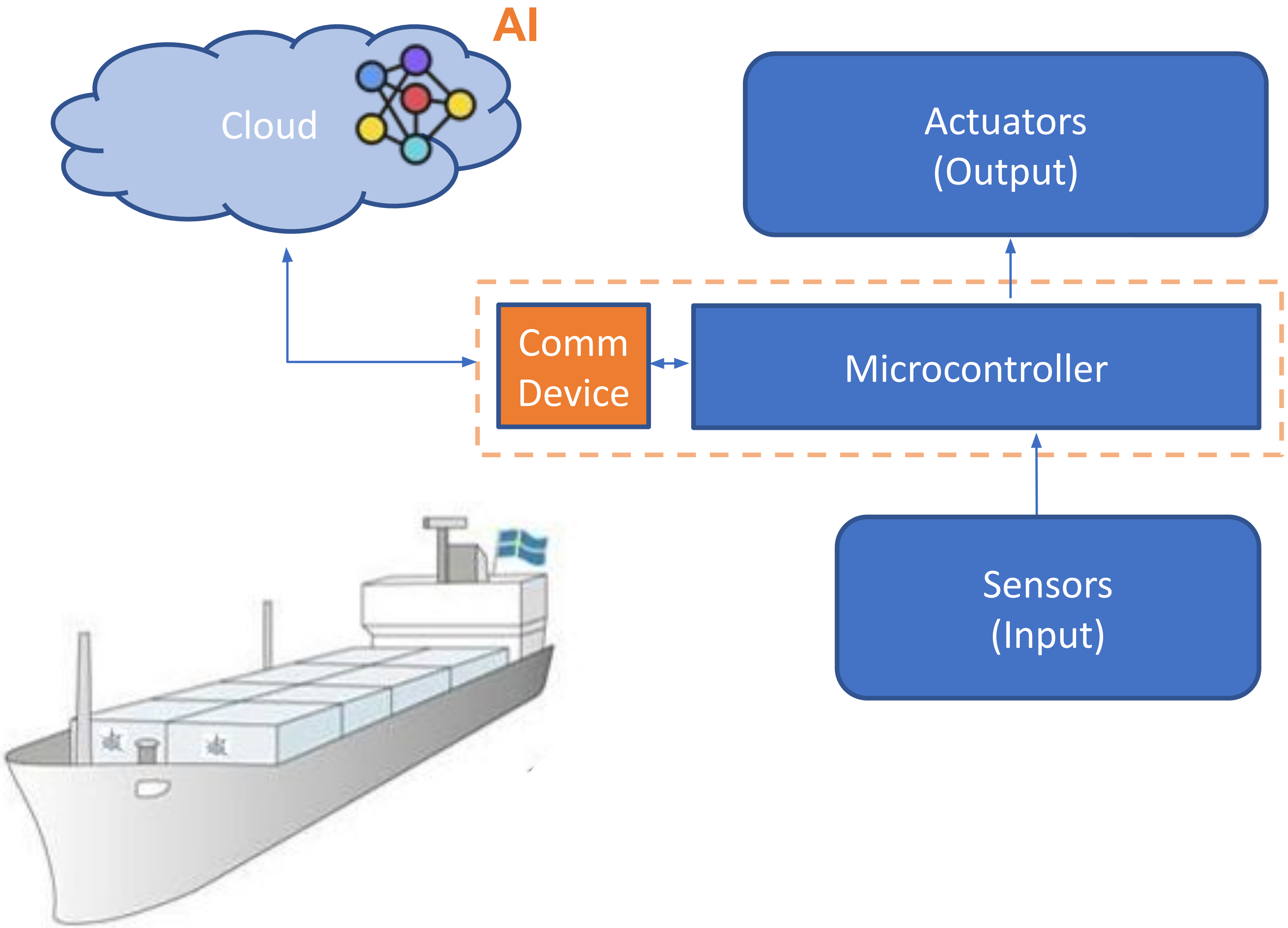
of unstructured data is analyzed or used at all

Source: Harvard Business Review, [What's Your Data Strategy?](), April 18, 2017
Cisco, [Internet of Things (IoT) Data Continues to Explode Exponentially. Who Is Using That Data and How?](), Feb 5, 2018

Typical AIoT Project

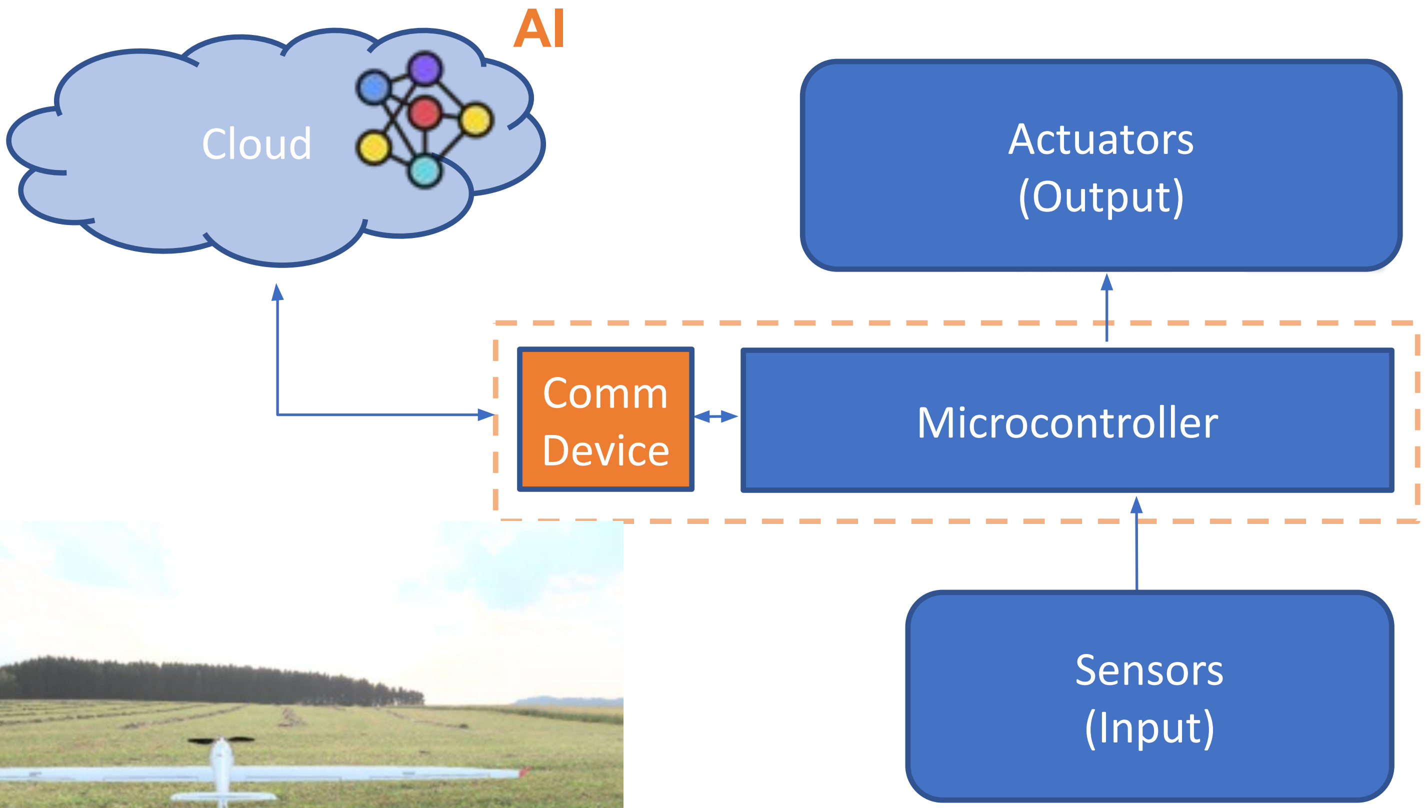


Typical AIoT Project … … Issues



Bandwidth

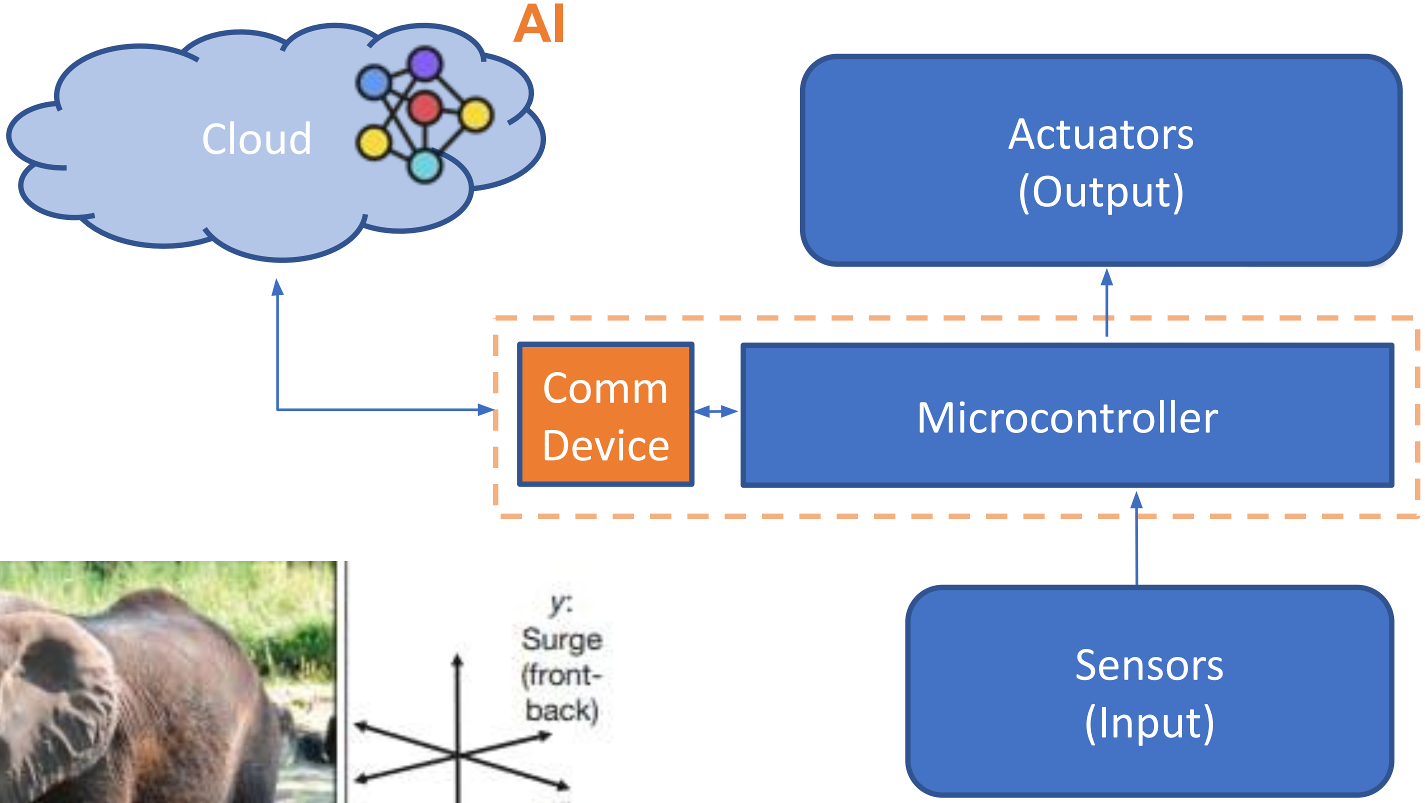
Typical AIoT Project … … Issues

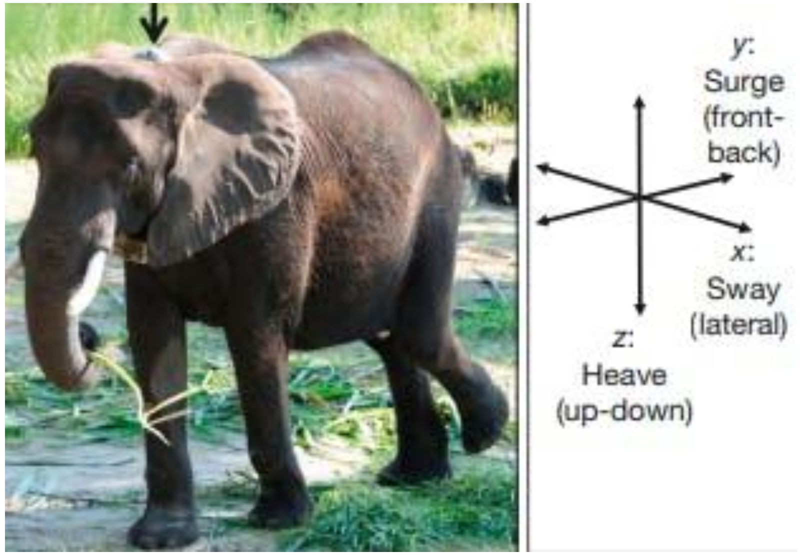


Bandwidth

Latency

Typical AIoT Project … … Issues



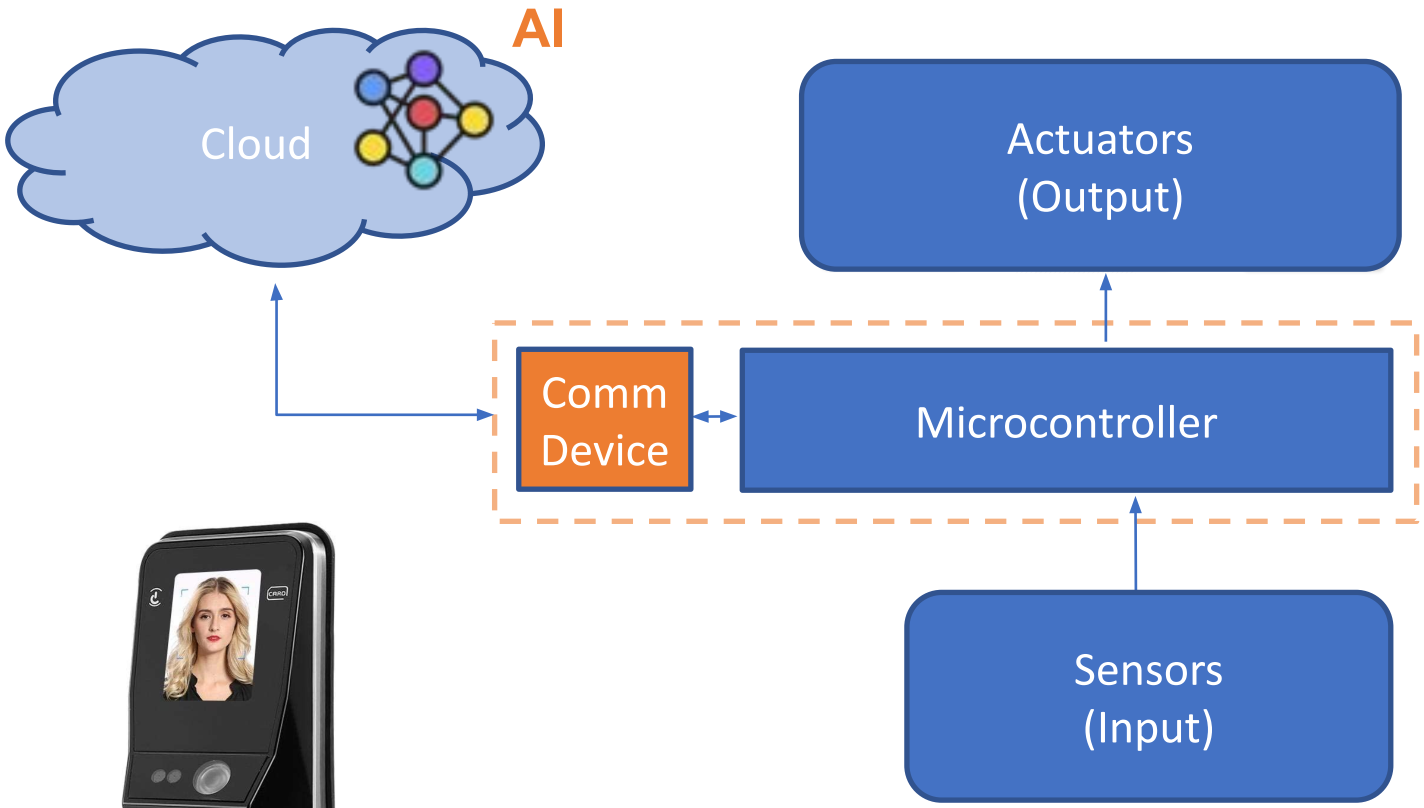


Bandwidth

Latency

Energy

Typical AIoT Project … … Issues



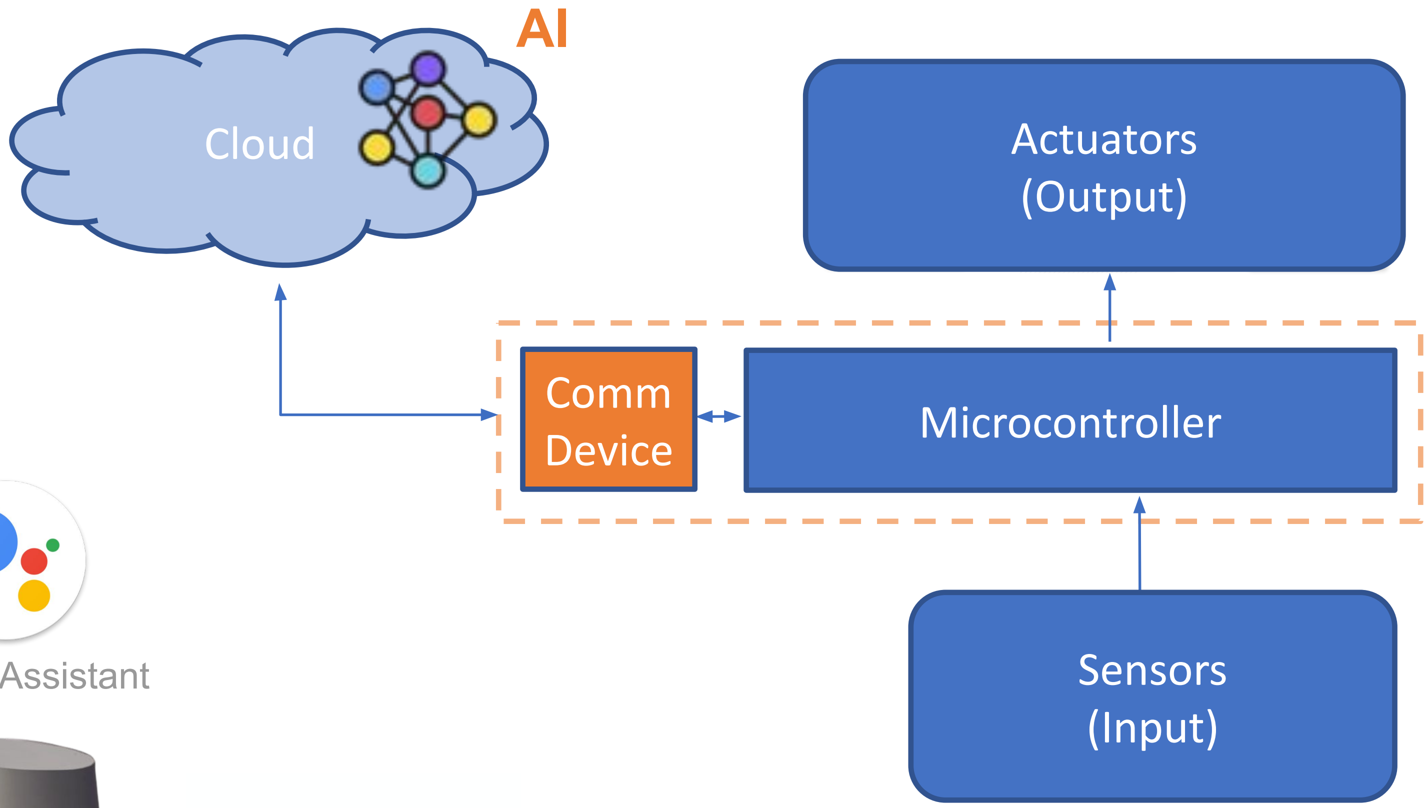
Bandwidth

Latency

Energy

Reliability

Typical AIoT Project … … Issues



Google Assistant

Bandwidth

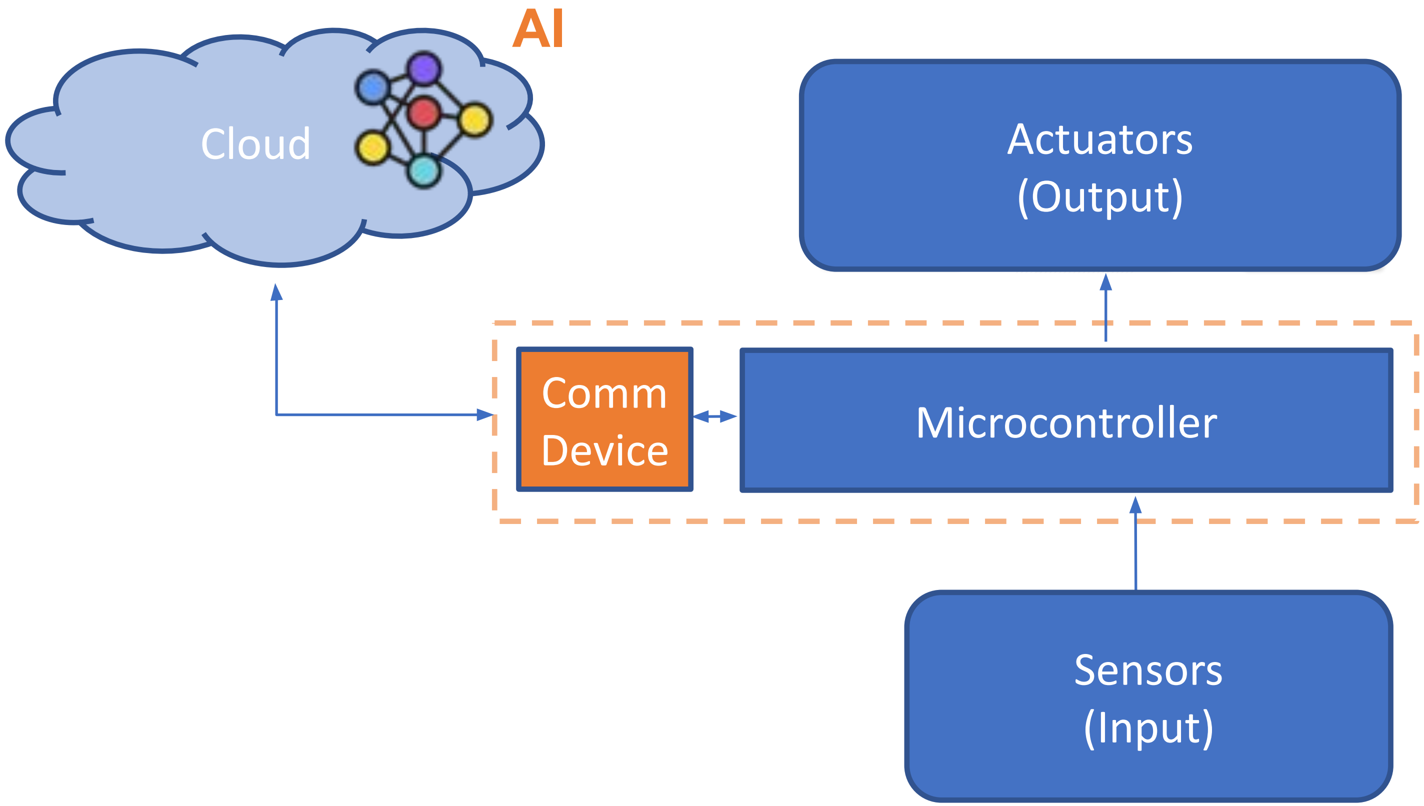
Latency

Energy

Reliability

Privacy

Typical AIoT Project … … Issues

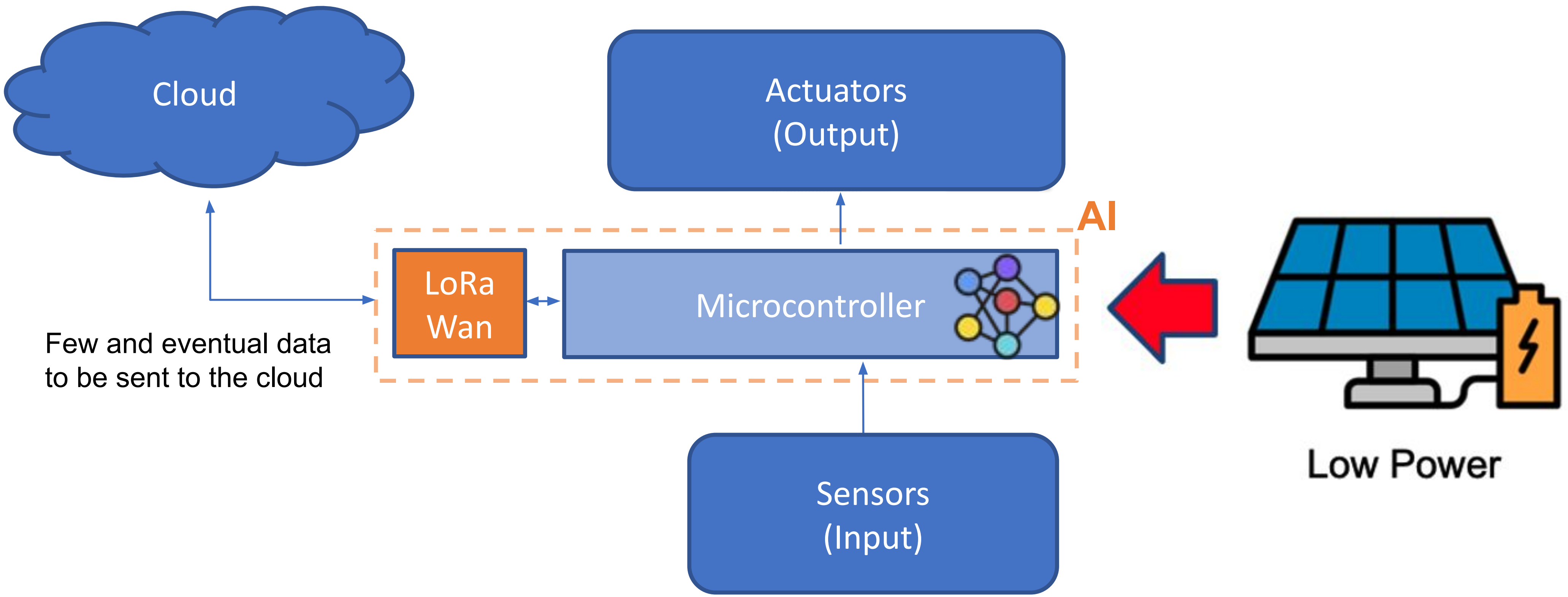


Bandwidth
Latency
Energy
Reliability
Privacy

… Solution ?

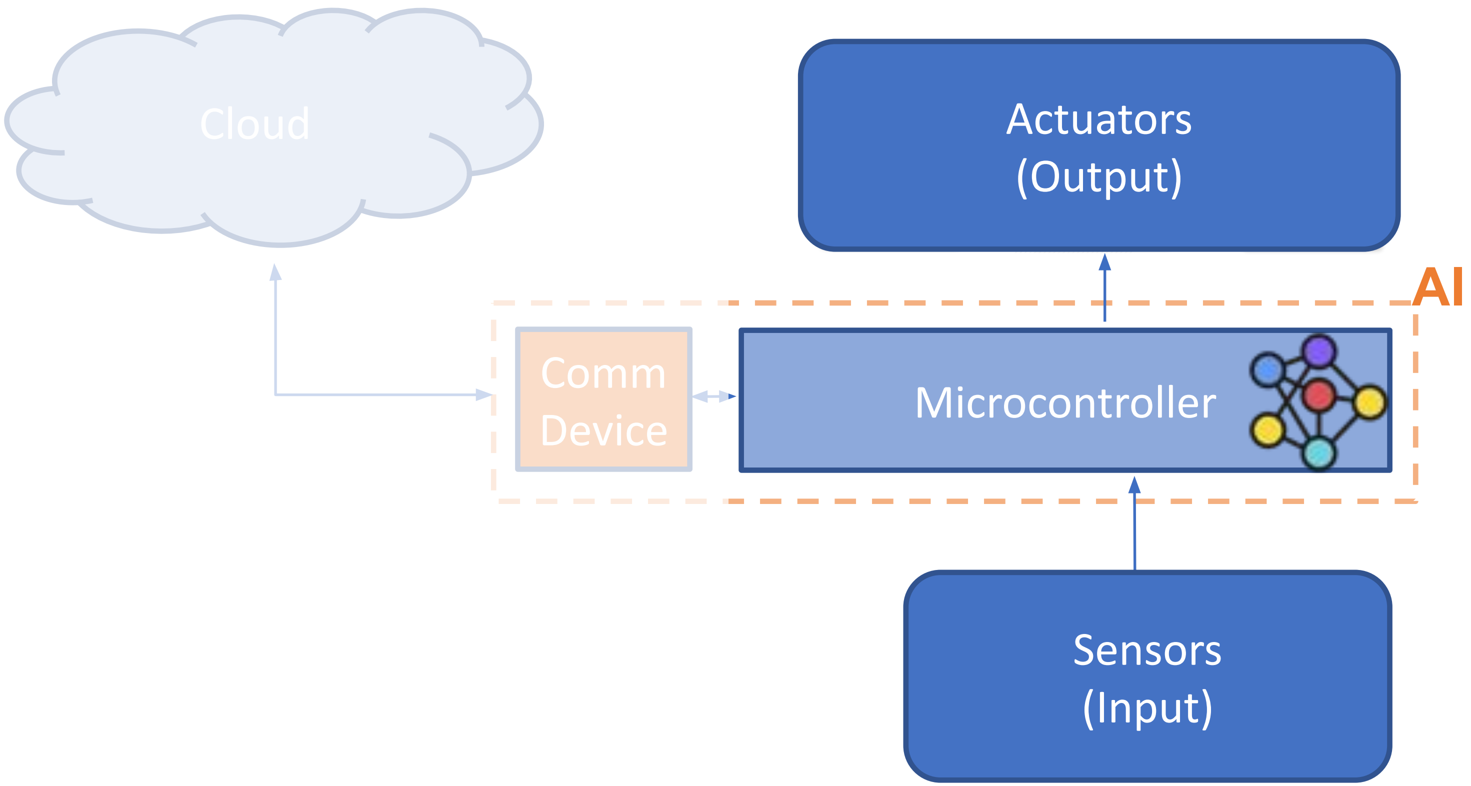
IoT 2.0 * – Edge AI/ML

* Intelligence of Things



... **Solution ->** ML goes close to data

When to use an Edge AI/ML approach:



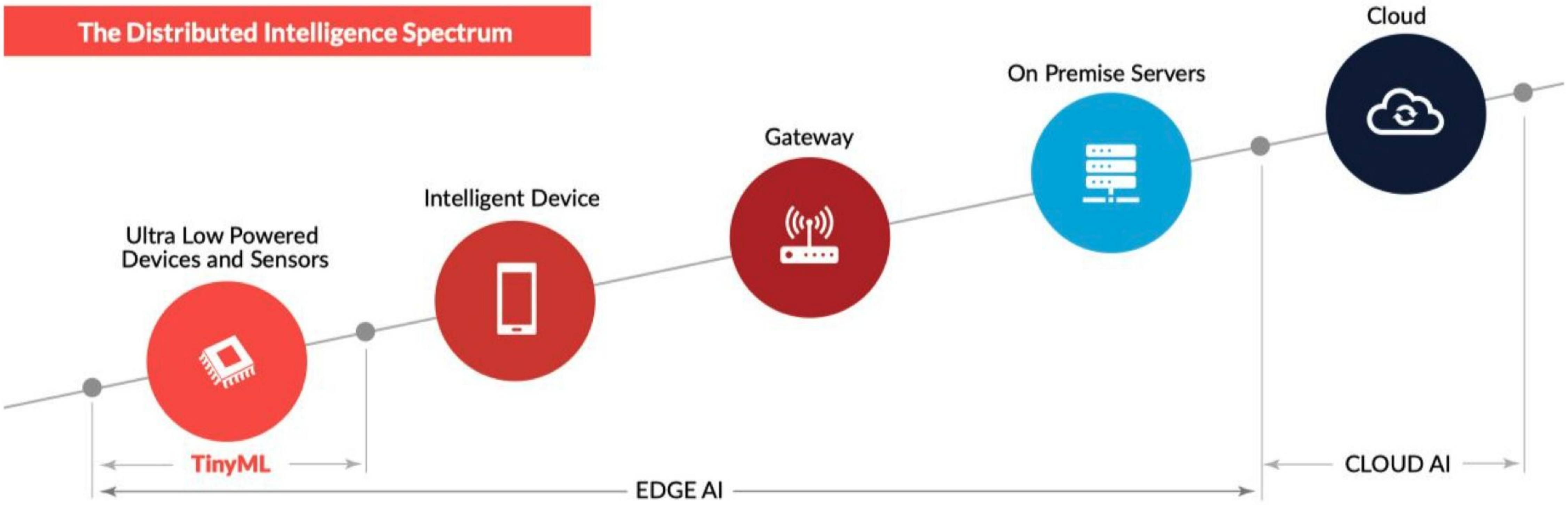
Bandwidth

Latency

Energy

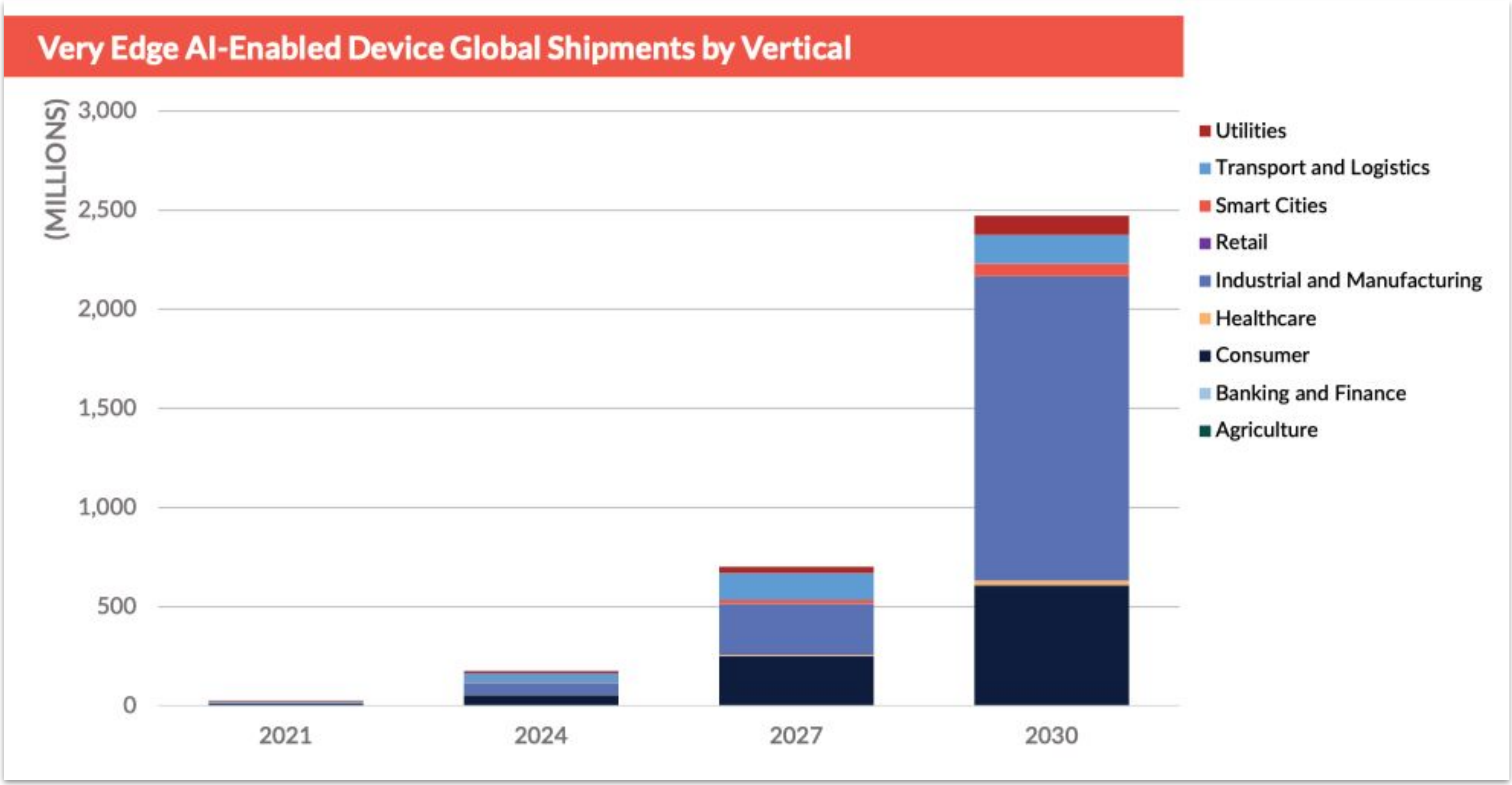
Reliability

Privacy



Source: ABI Research: TinyML

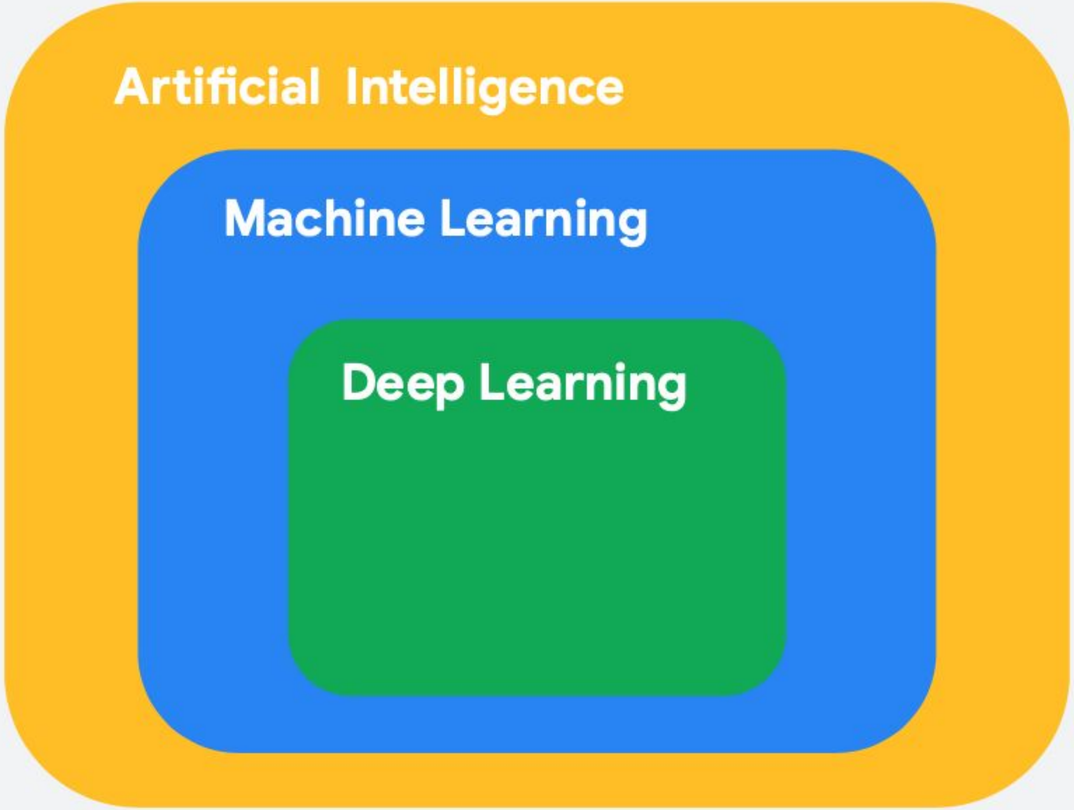
Market Forecast



Source: ABI Research: TinyML

Embedded ML (TinyML)

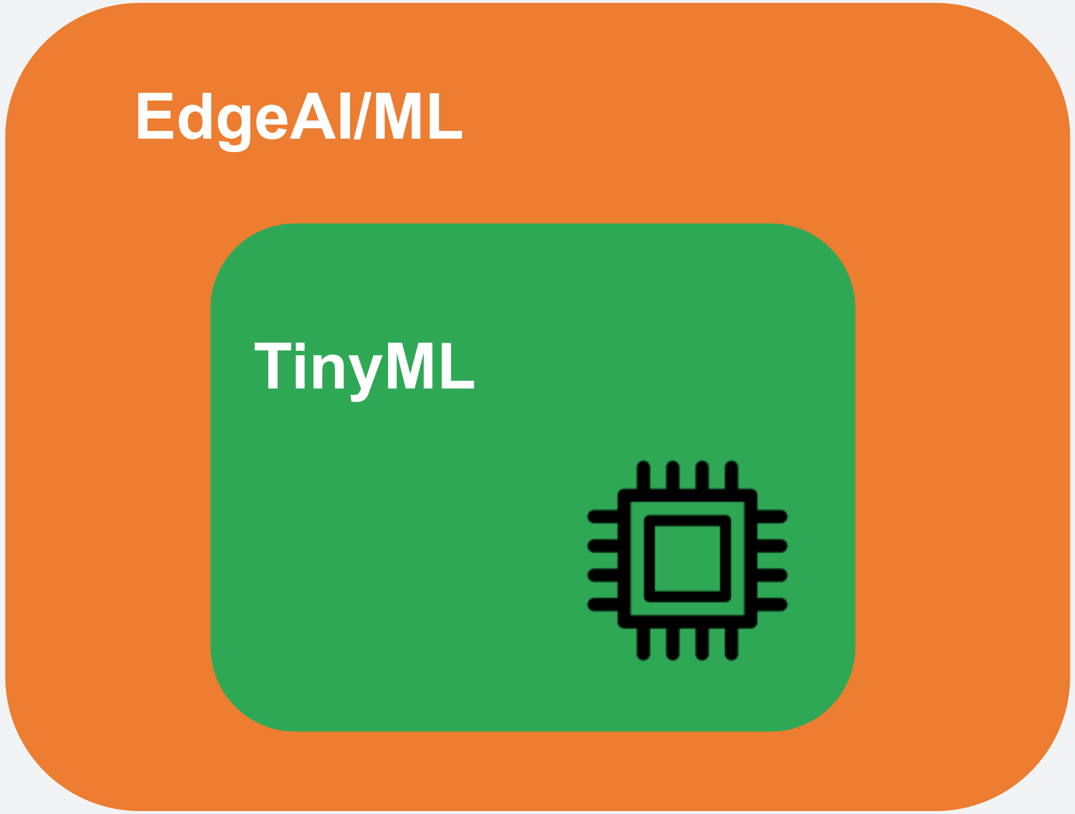
Introduction



AI: Any technique that enables computers to mimic human behavior

ML: Ability to learn without explicitly being programed

DL: Extract patterns from data using neural networks



Edge AI (or Edge ML) is the processing of Artificial Intelligence algorithms on edge, that is, on users' devices. The concept derives from **Edge Computing**, which starts from the same premise: data is stored, processed, and managed directly at the Internet of Things (IoT) endpoints.

TinyML is a subset of EdgeML, where sensors are generating data with ultra-low power consumption (batteries), so that we can ultimately deploy machine learning continuously ("always on devices")

What is Tiny Machine Learning (TinyML)?

TinyML

Fastest-growing field of **ML**

20

What is Tiny Machine Learning (**TinyML**)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

21

What is Tiny Machine Learning (TinyML)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

On-device sensor analytics

22

What is Tiny Machine Learning (TinyML)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

On-device sensor analytics

Low power consumption

23

What is Tiny Machine Learning (**TinyML**)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

On-device sensor analytics

Low power consumption

Battery-operated

24

What is Tiny Machine Learning (**TinyML**)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

On-device sensor analytics

Low power consumption

Battery-operated

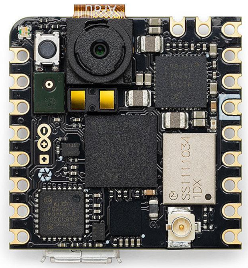
Always-on ML

25

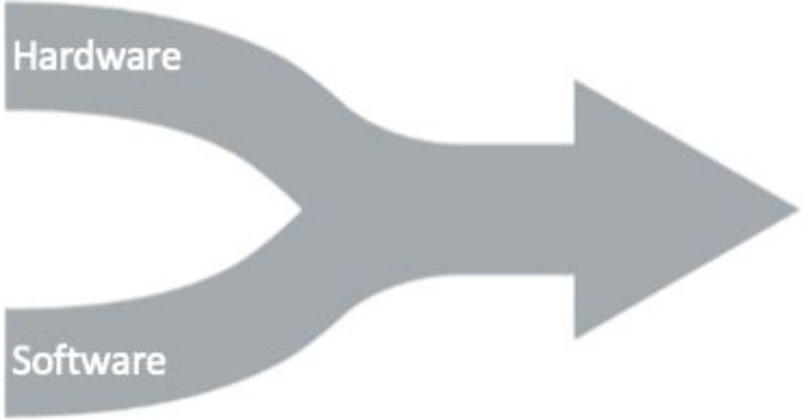
What Makes TinyML ?

TinyML

What Makes **TinyML** ?

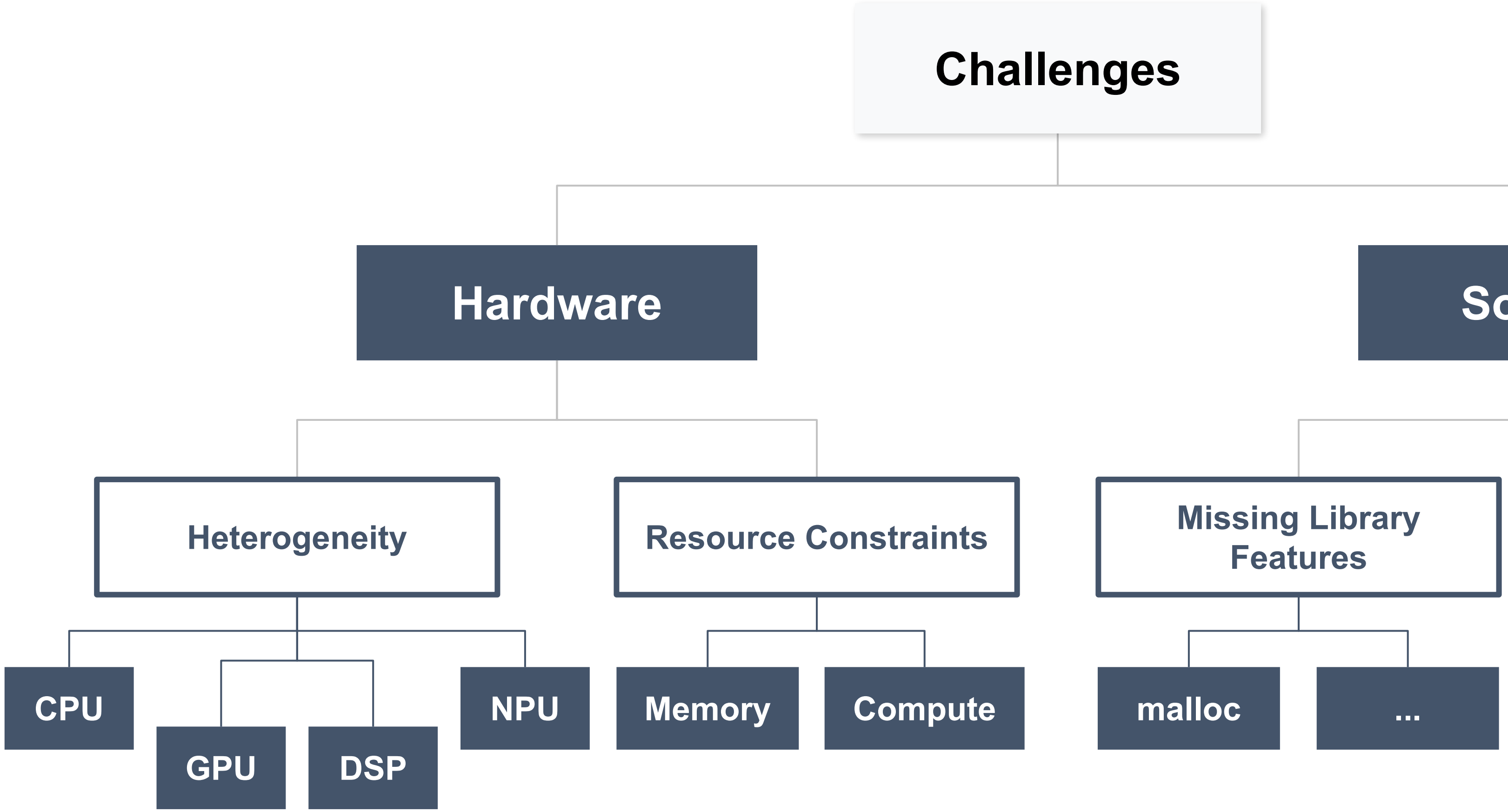


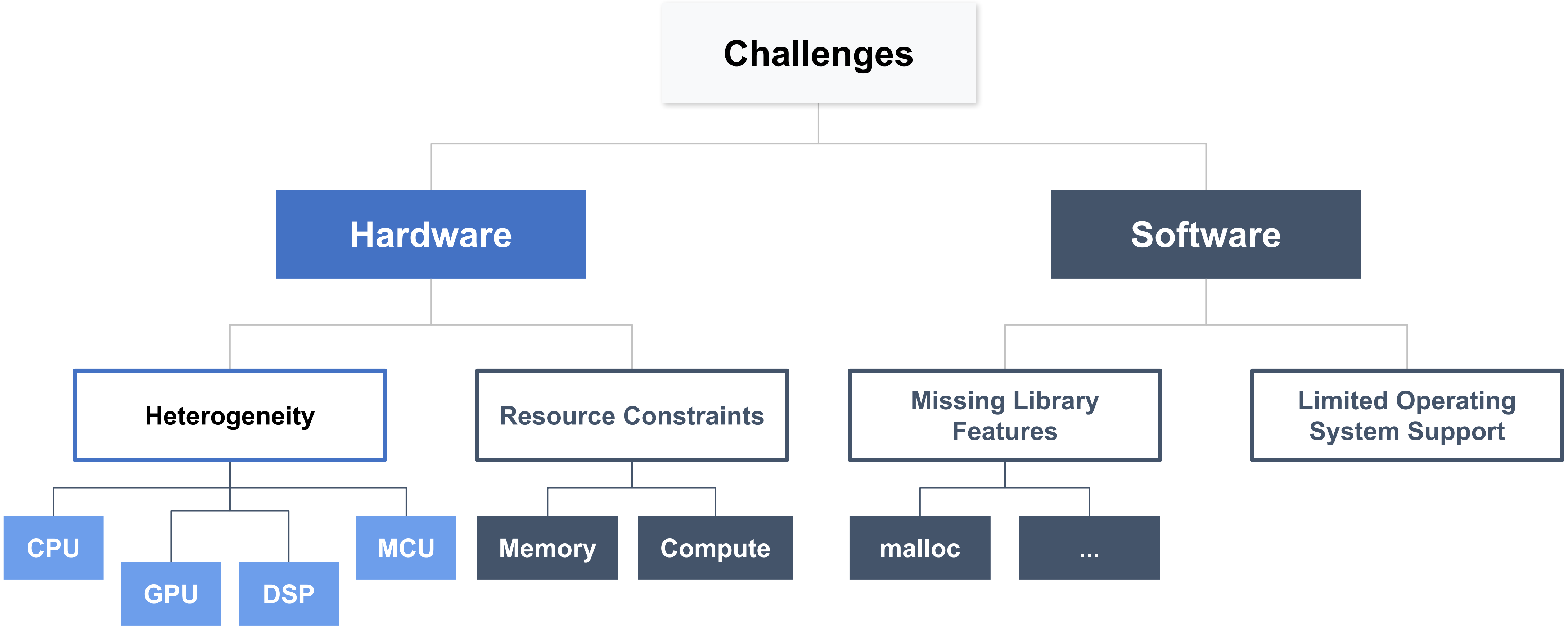
TensorFlow Lite



TinyML

TinyML Challenges


Challenges
Hardware
Software
Heterogeneity
Resource Constraints
Missing Library Features
Limited Operating System Support
CPU
GPU
DSP
NPU
Memory
Compute
malloc
...

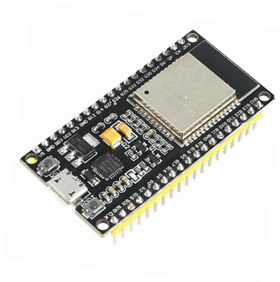

Challenges
Hardware
Software
Heterogeneity
Resource Constraints
Missing Library Features
Limited Operating System Support
CPU
GPU
DSP
MCU
Memory
Compute
malloc
...

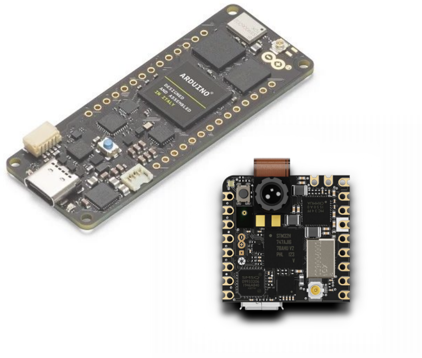
250 Billion

MCUs today

31

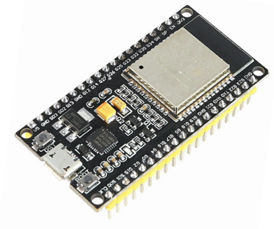
Hardware

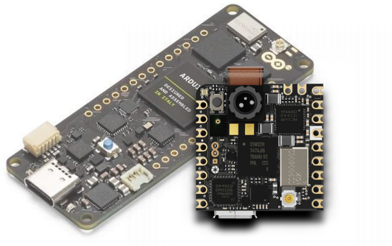




Hardware

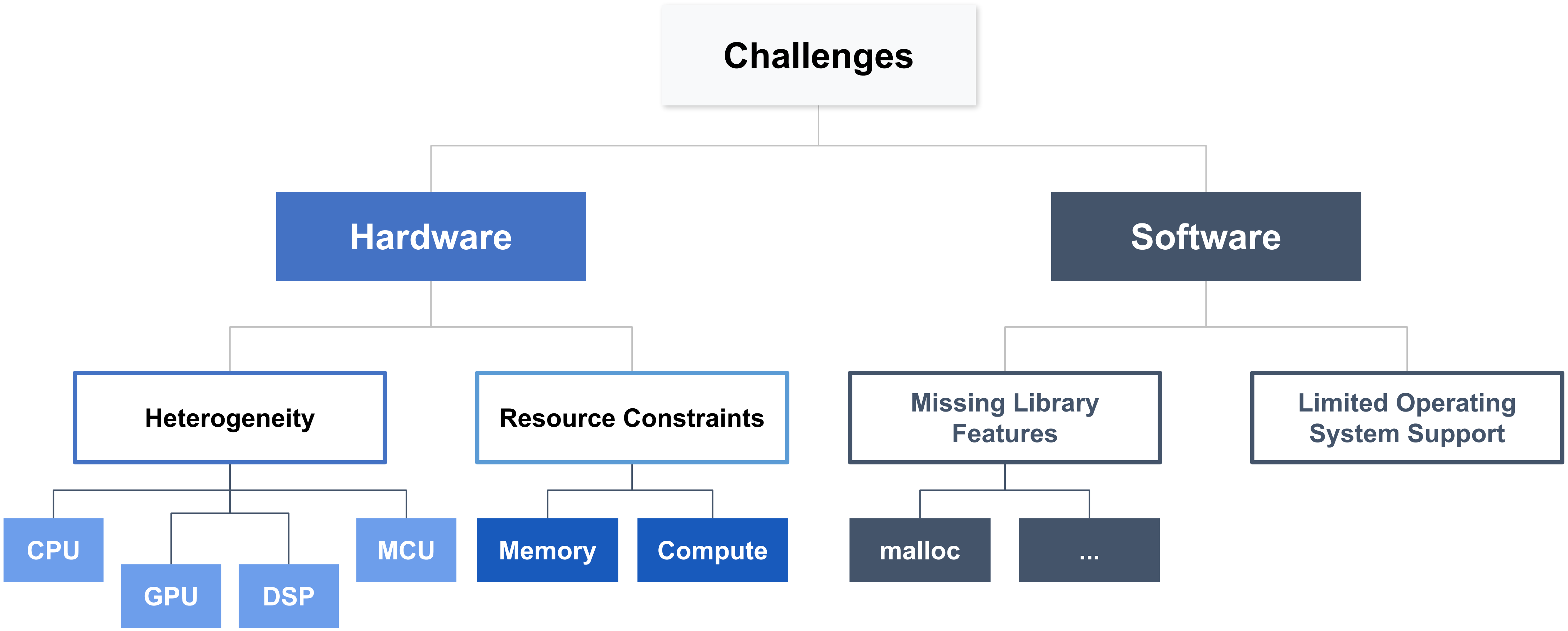




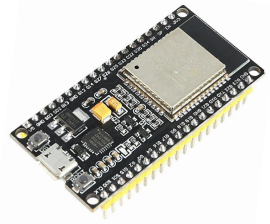
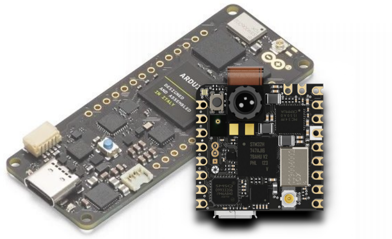


	Raspberry Pico (W)	Arduino Nano Sense	ESP 32	Seeed XIAO Sense / ESP32S3	Arduino Pro
32Bits CPU	Dual-core Arm Cortex-M0+	Arm Cortex-M4F	Xtensa LX6 Dual Core	Arm Cortex-M4F (BLE) Xtensa LX7 Dual Core	Dual Core Arm Cortex M7/M4
CLOCK	133MHz	64MHz	240MHz	64 / 240MHz	480/240MHz
RAM	264KB	256KB	520KB (part available)	256KB / 8MB	1MB
ROM	2MB	1MB	2MB	2MB / 8MB	2MB
Radio	(Yes for W)	BLE	BLE/WiFi	BLE / WiFi (ESP32S3)	BLE/WiFi
Sensors	No	Yes	No	Yes (Sense)	Yes (Nicla)
Bat. Power Manag.	No	No	No	Yes	Yes
Price	$	$$$	$	$$	$$$$$

https://media.digikey.com/Resources/Maker/the-original-guide-to-boards-2022.pdf

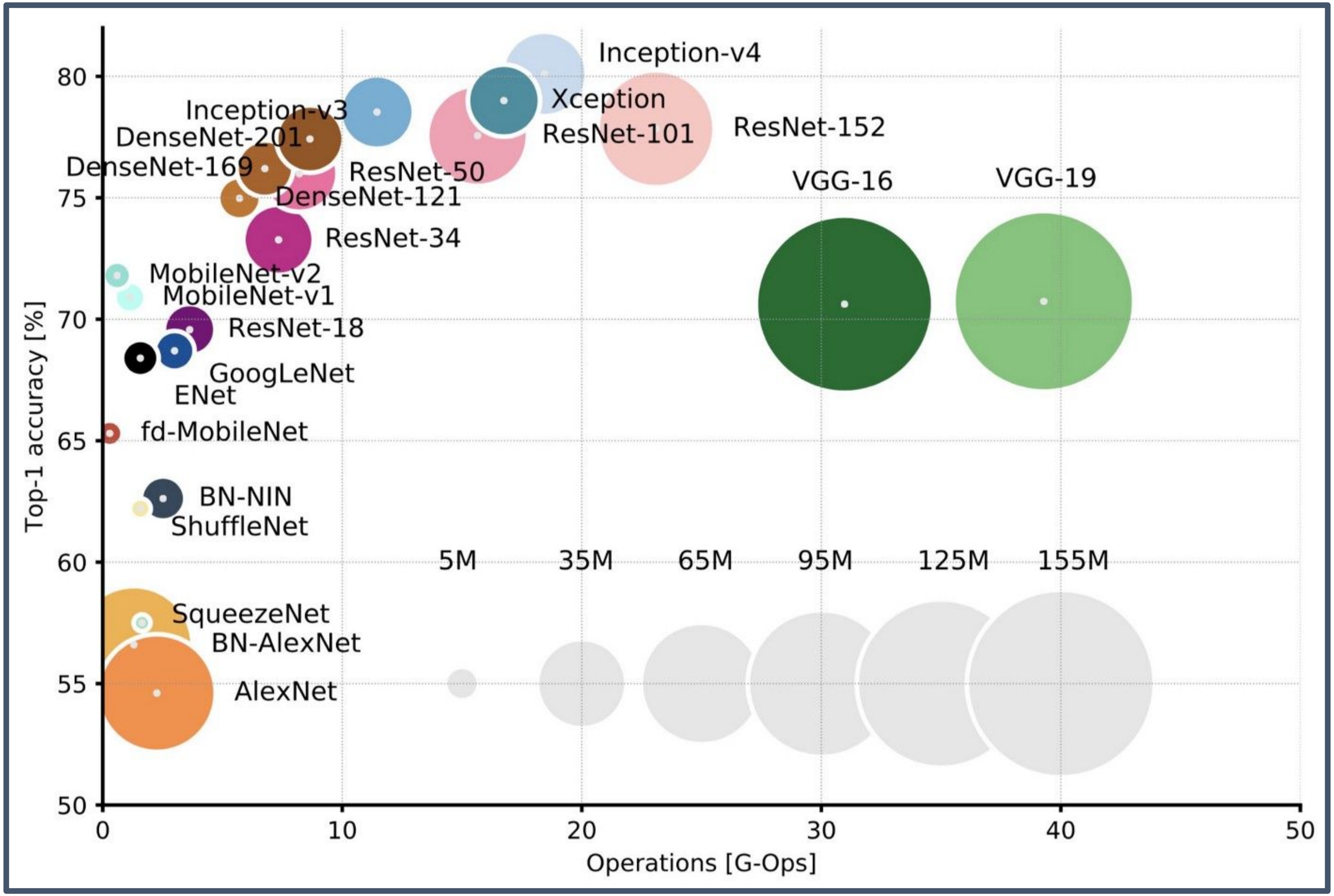

Challenges
Hardware
Software
Heterogeneity
Resource Constraints
Missing Library Features
Limited Operating System Support
CPU
GPU
DSP
MCU
Memory
Compute
malloc
...

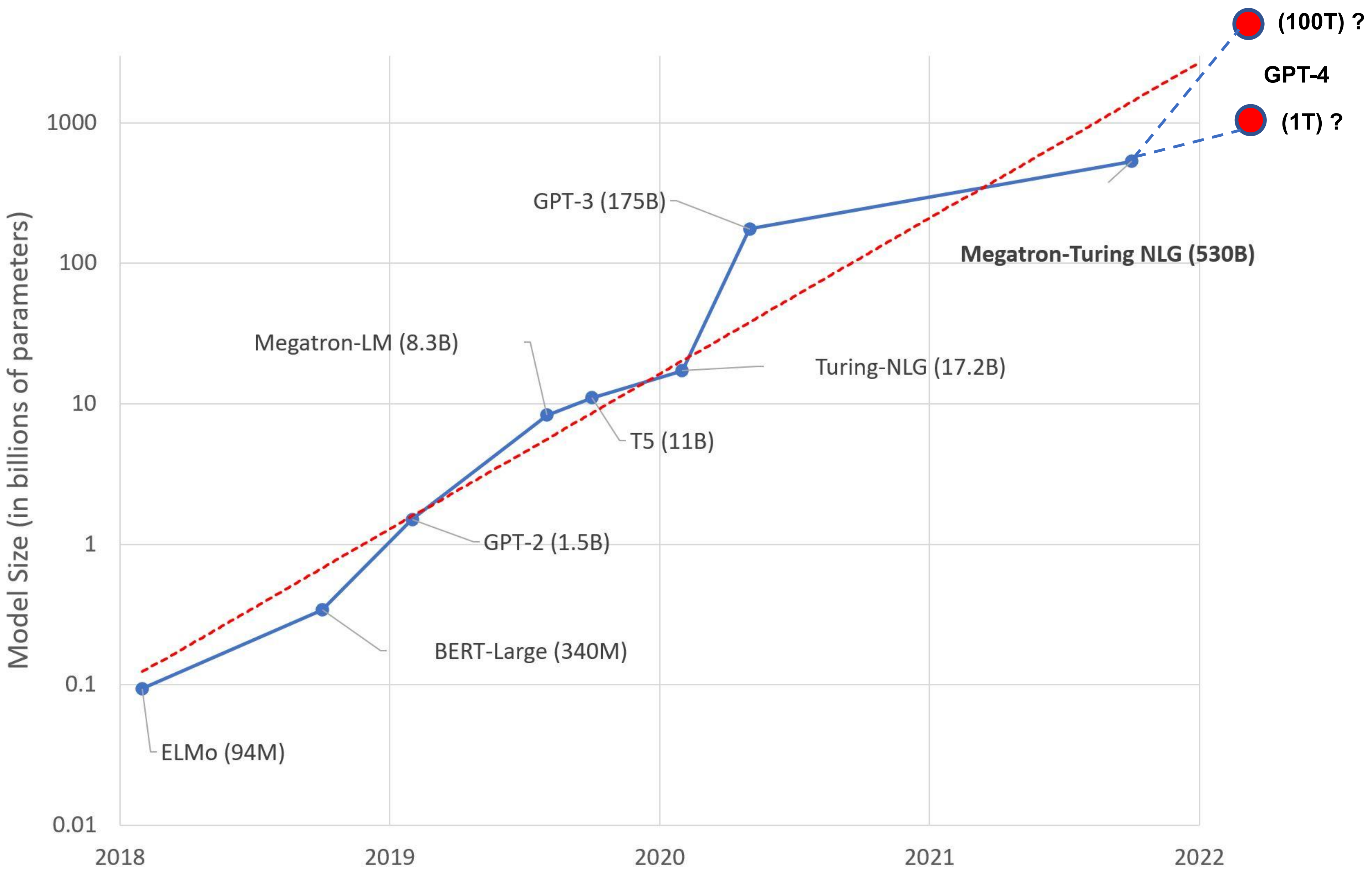
Hardware

					
	Raspberry Pico (W)	Arduino Nano Sense	ESP 32	Seeed XIAO Sense / ESP32S3	Arduino Pro
32Bits CPU	Dual-core Arm Cortex-M0+	Arm Cortex-M4F	Xtensa LX6 Dual Core	Arm Cortex-M4F (BLE) Xtensa LX7 Dual Core	Dual Core Arm Cortex M7/M4
CLOCK	133MHz	64MHz	240MHz	64 / 240MHz	480/240MHz
RAM	264KB	256KB	520KB (part available)	256KB / 8MB	1MB
ROM	2MB	1MB	2MB	2MB / 8MB	2MB
Radio	(Yes for W)	BLE	BLE/WiFi	BLE / WiFi (ESP32S3)	BLE/WiFi
Sensors	No	Yes	No	Yes (Sense)	Yes (Nicla)
Bat. Power Manag.	No	No	No	Yes	Yes
Price	$	$$$	$	$$	$$$$$

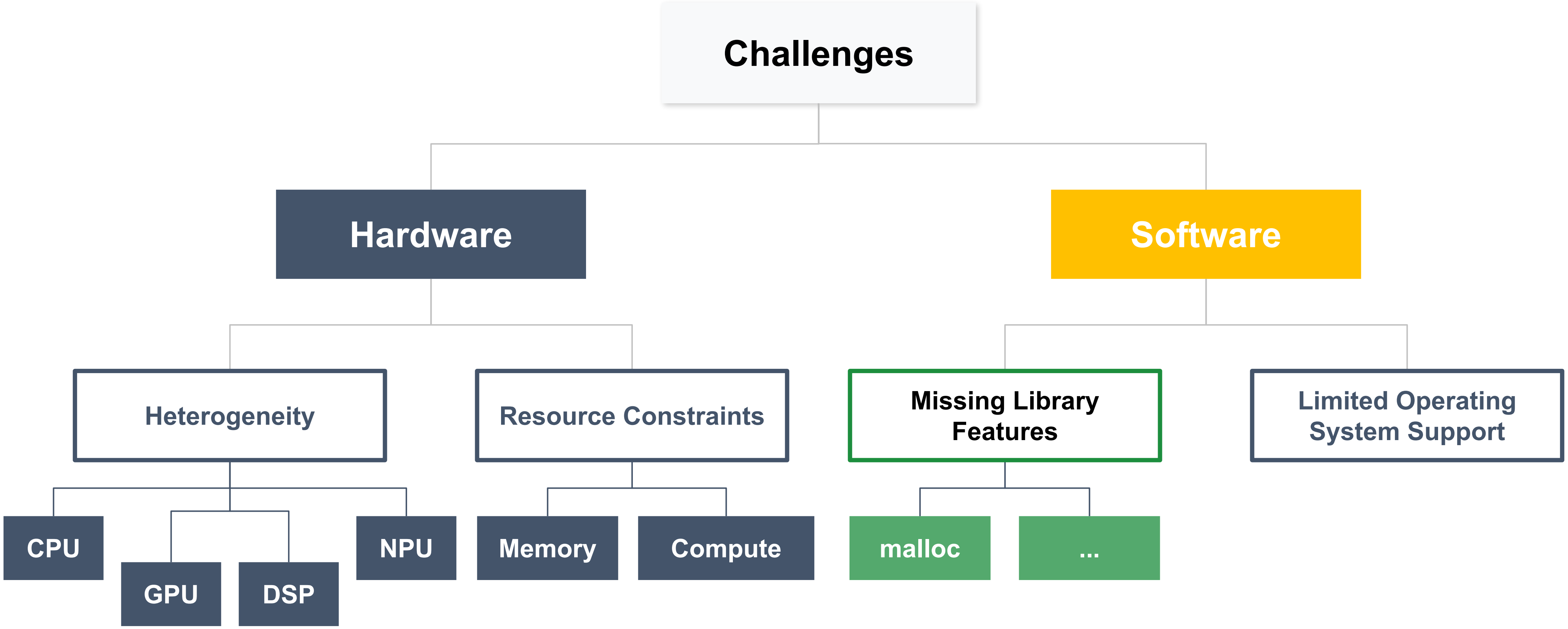
https://media.digikey.com/Resources/Maker/the-original-guide-to-boards-2022.pdf

https://arxiv.org/pdf/1910.01108.pdf
https://towardsdatascience.com/neural-network-architectures-156e5bad51ba

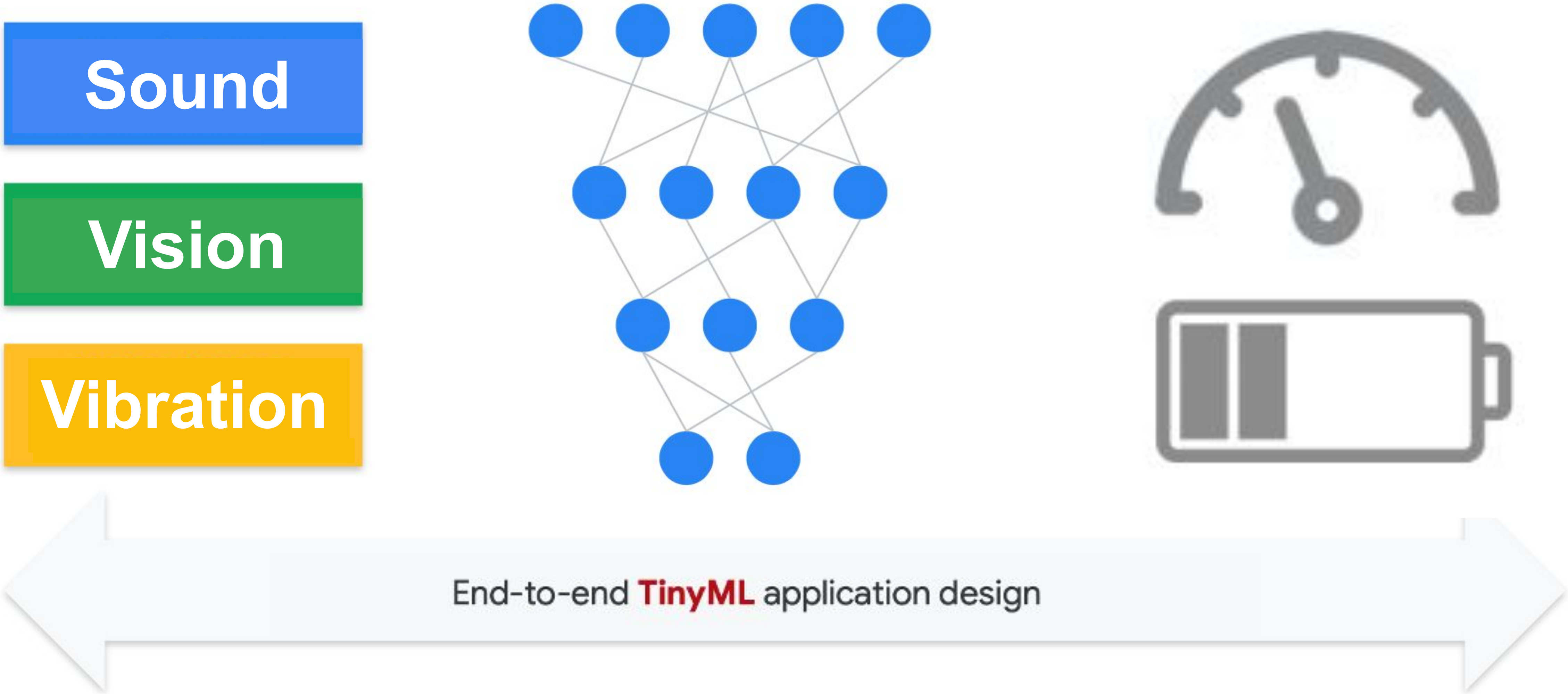




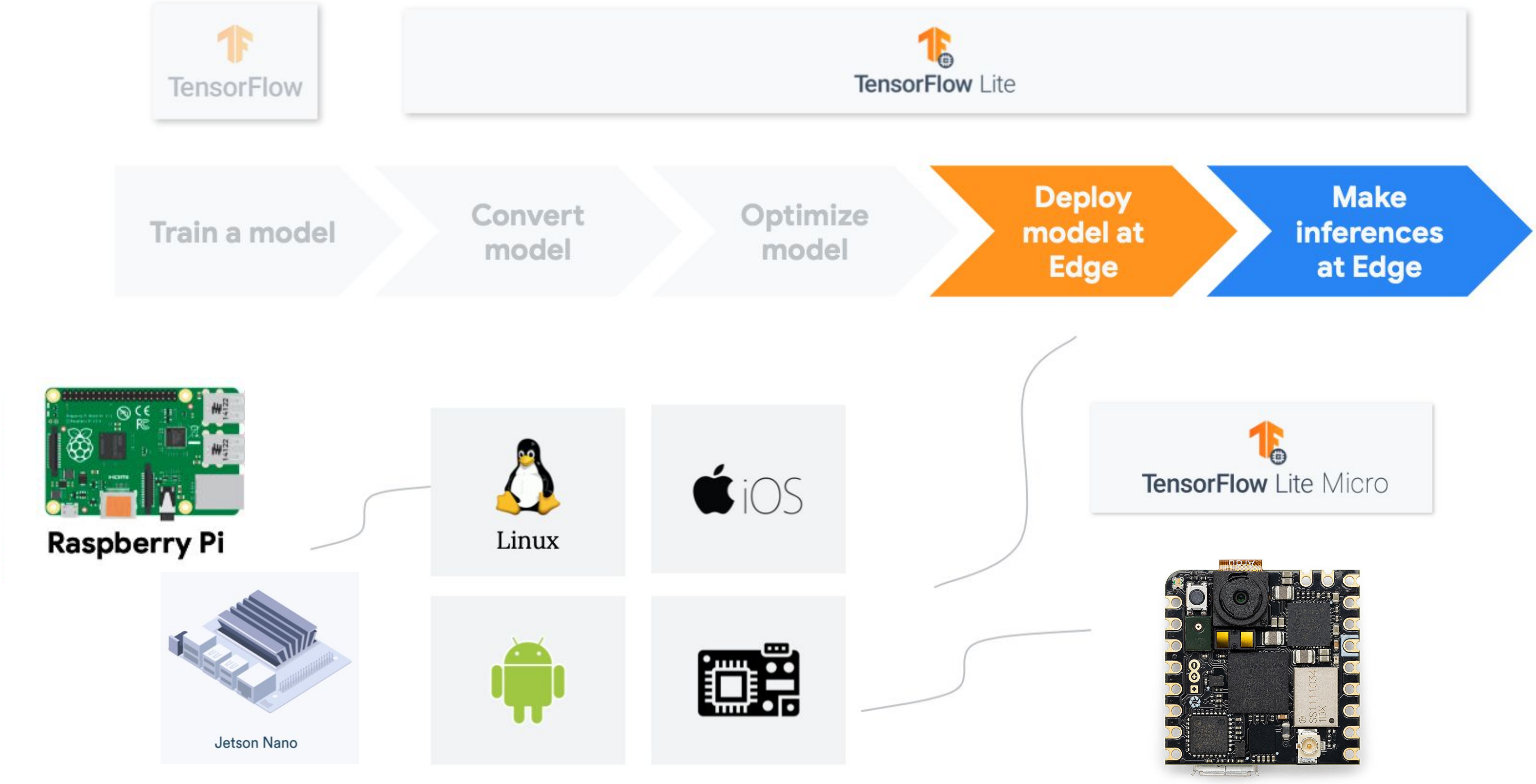
https://huggingface.co/blog/large-language-models https://arxiv.org/abs/2303.08774

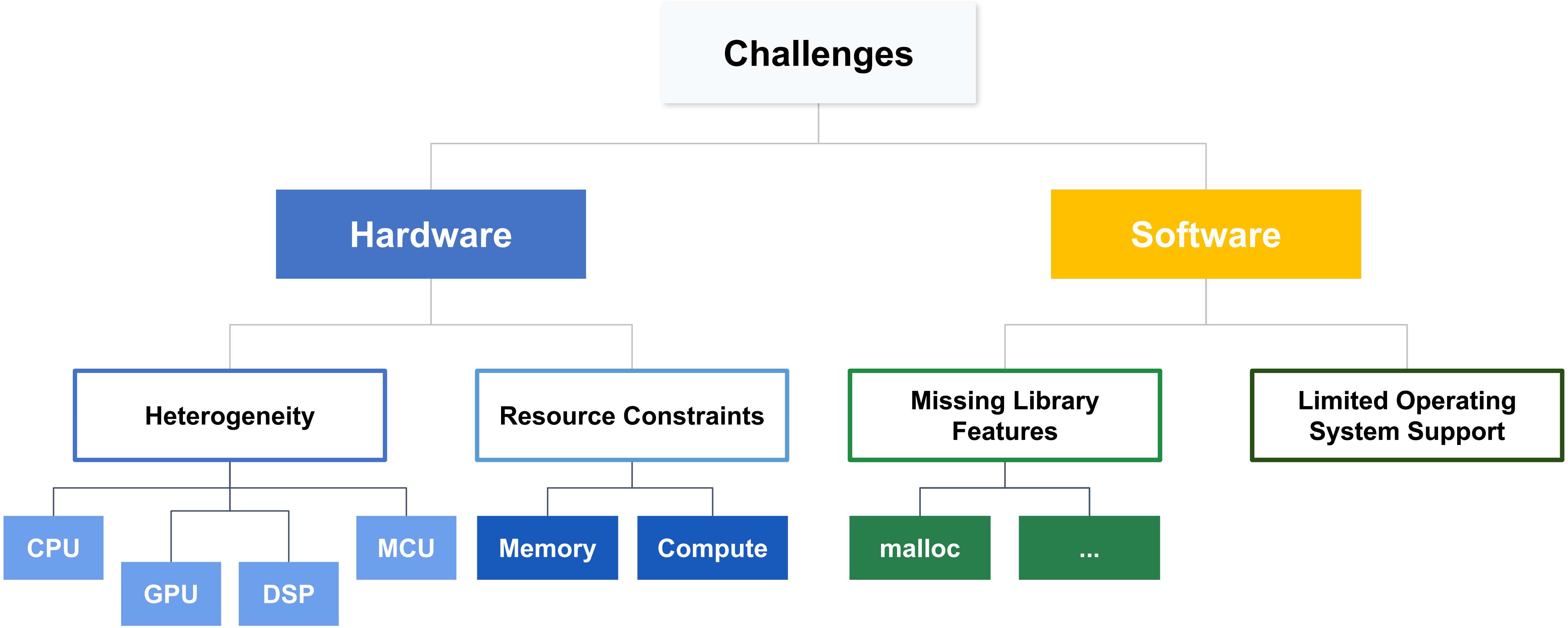

Challenges
Hardware
Software
Heterogeneity
Resource Constraints
Missing Library Features
Limited Operating System Support
CPU
GPU
DSP
NPU
Memory
Compute
malloc
...

Datasets Preprocessing | **Quantization Pruning** | **Resource constraints**

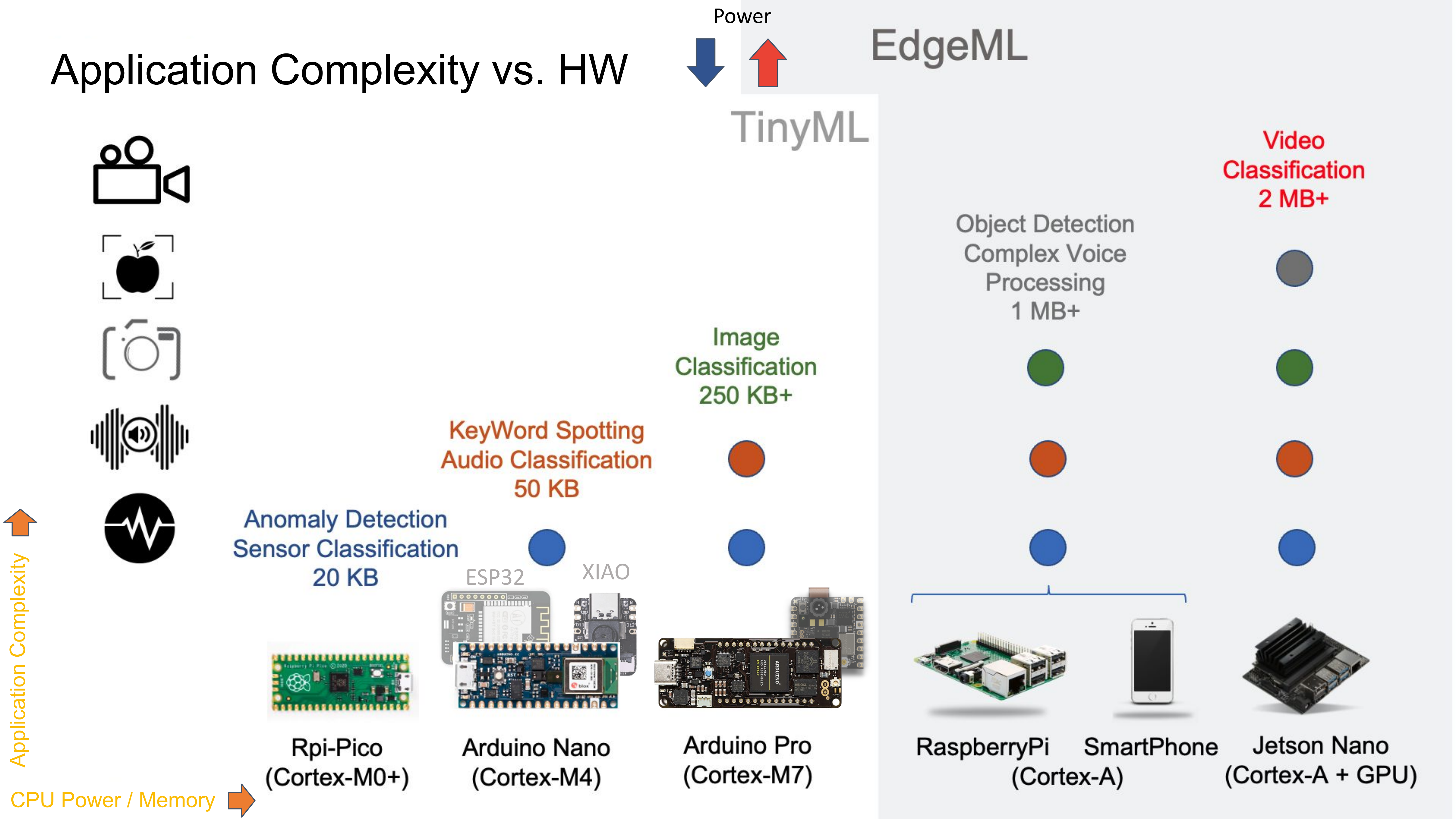


Software


TensorFlow
TensorFlow Lite
Train a model
Convert model
Optimize model
Deploy model at Edge
Make inferences at Edge
Raspberry Pi
Jetson Nano
Linux
iOS
TensorFlow Lite Micro

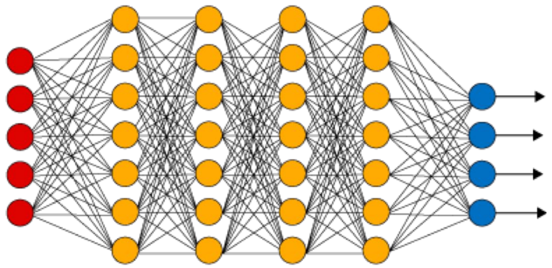

Challenges
Hardware
Software
Heterogeneity
Resource Constraints
Missing Library Features
Limited Operating System Support
CPU
GPU
DSP
MCU
Memory
Compute
malloc
...

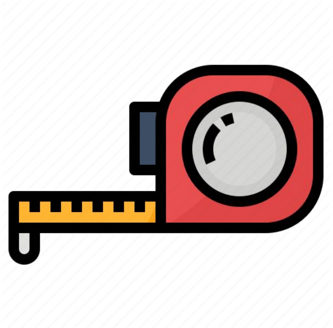
Application Complexity vs. HW



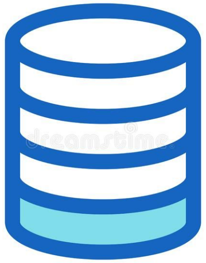
How to Train a ML Model?

Machine Learning Workflow ("What")

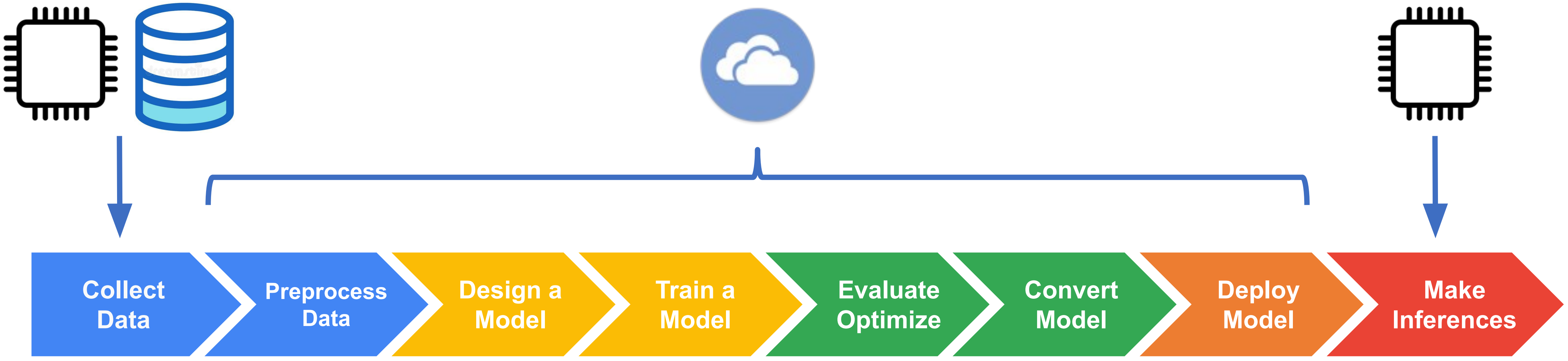




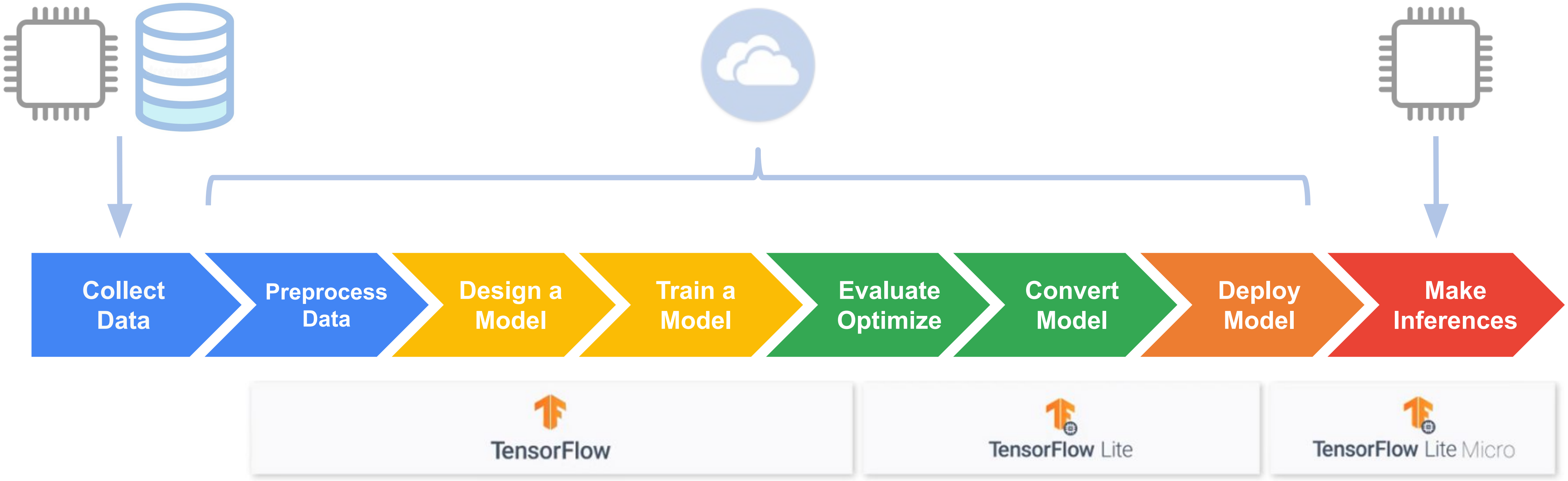
1. Collect Data
2. Preprocess Data
3. Design a Model
4. Train a Model
5. Evaluate Optimize
6. Convert Model
7. Deploy Model
8. Make Inferences



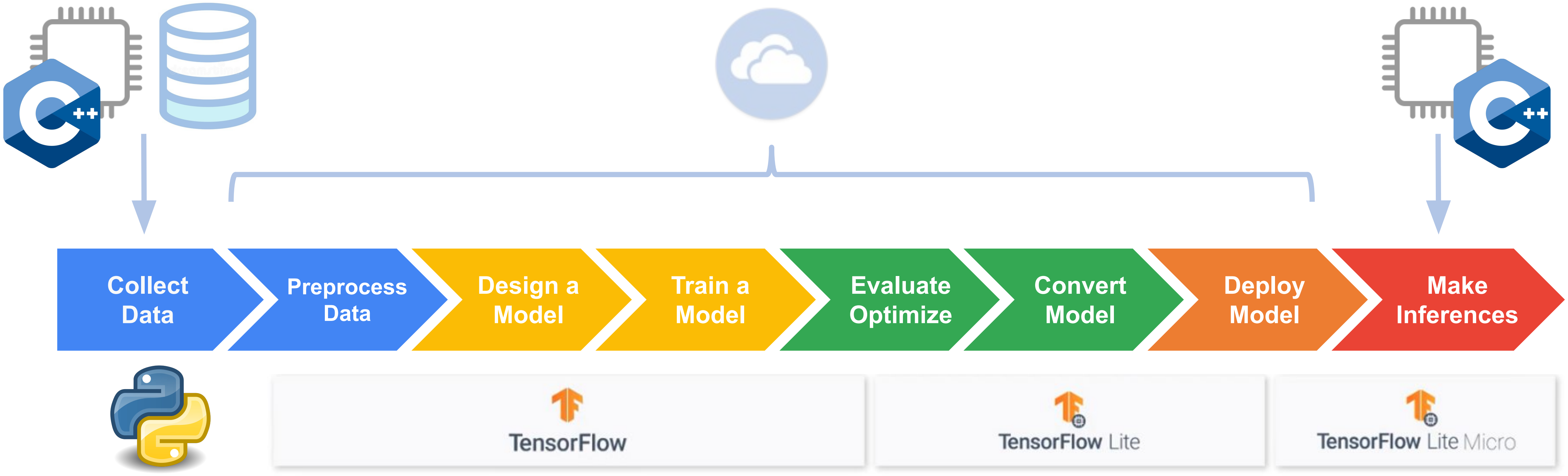
Machine Learning Workflow ("Where")



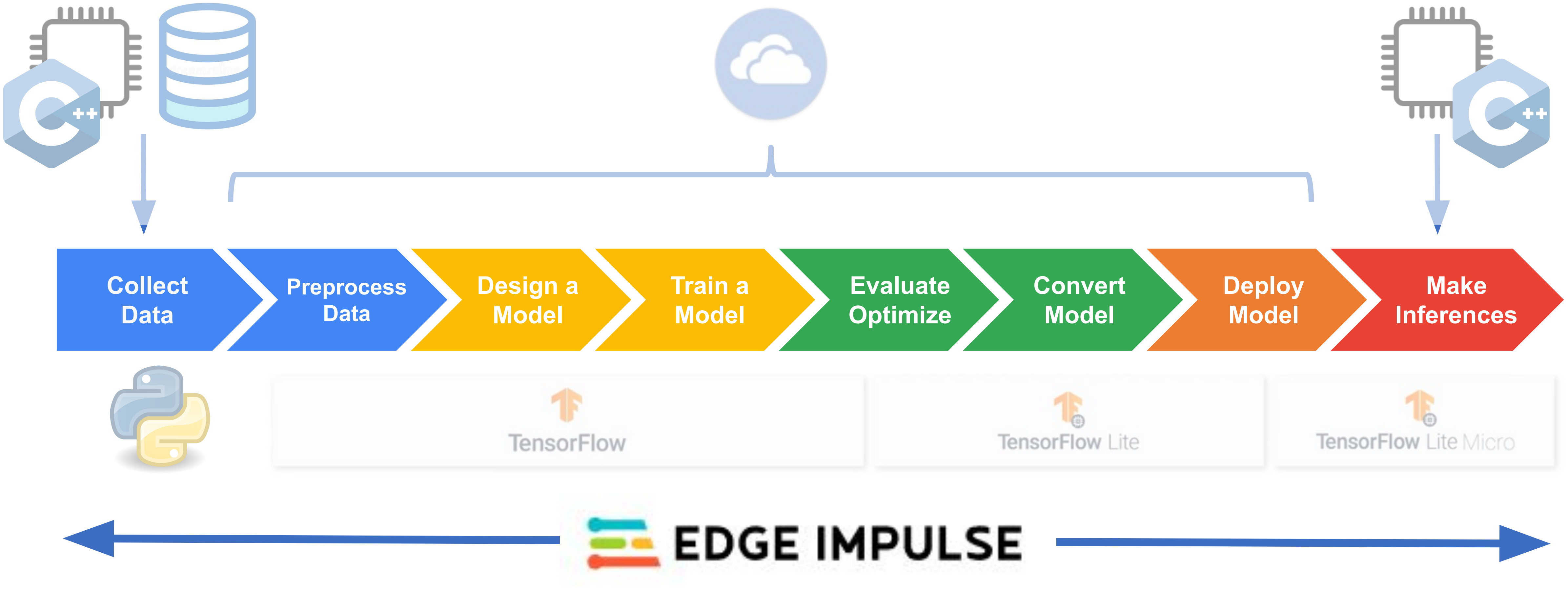
Machine Learning Workflow ("How")

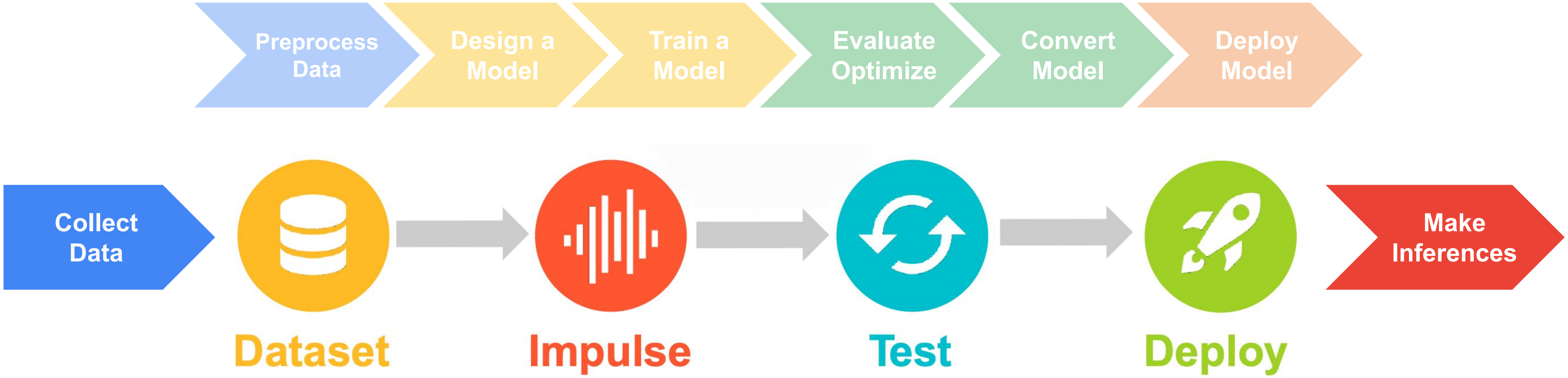


Machine Learning Workflow ("How")

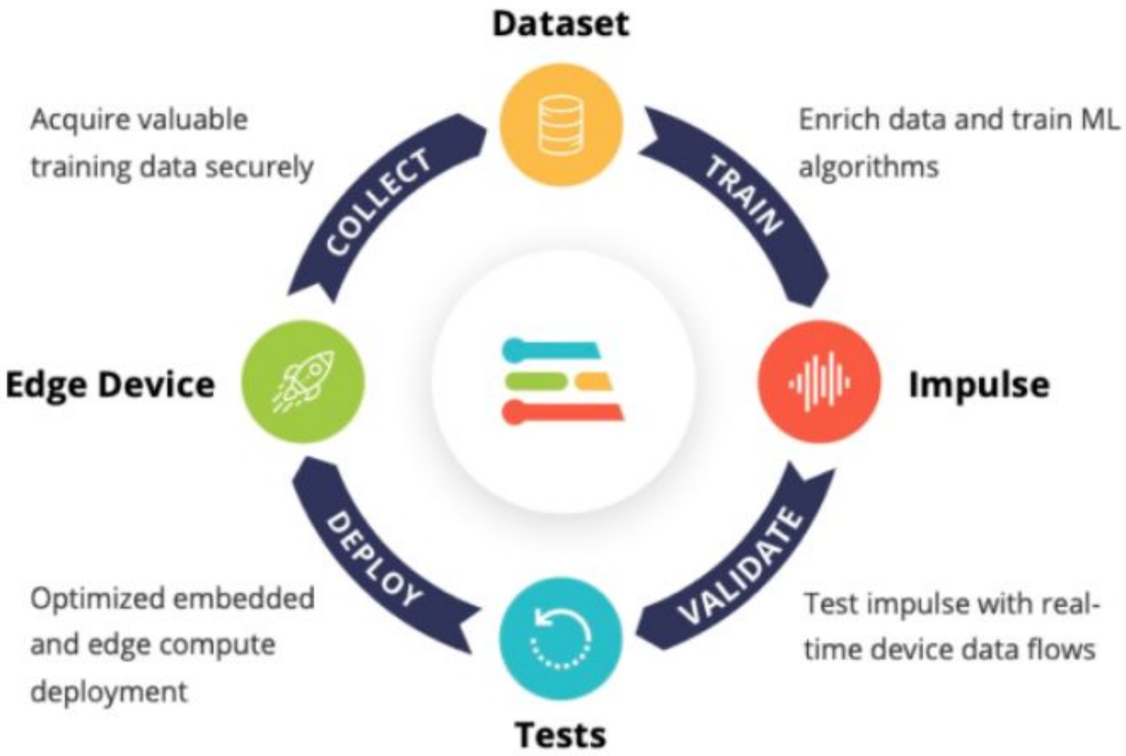


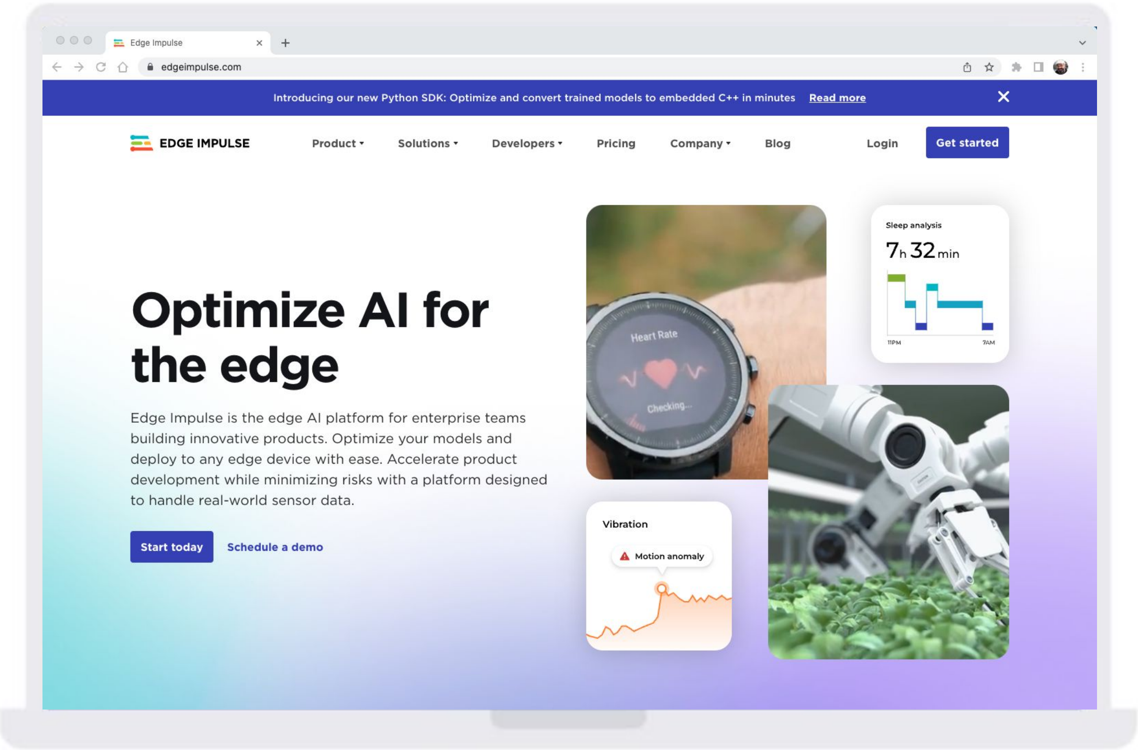
Machine Learning Workflow ("How")




Preprocess Data
Design a Model
Train a Model
Evaluate Optimize
Convert Model
Deploy Model
Collect Data
Dataset
Impulse
Test
Deploy
Make Inferences

EI Studio - Embedded ML platform (“AutoML”)



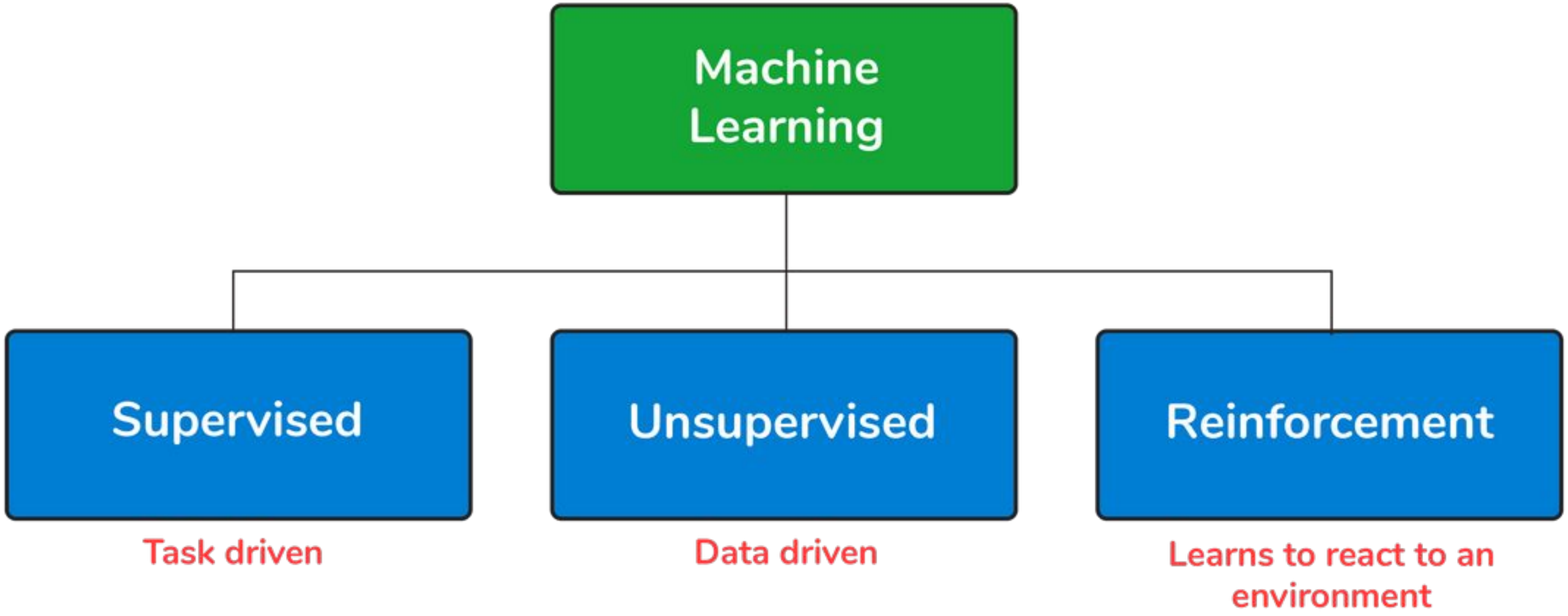


Learn more at http://edgeimpulse.com

[2023 SciTinyML Workshop - Edge Impulse Overview by Shawn Hymel]()

TinyML Application

Examples



- Regression
- Classification
- Object Detection

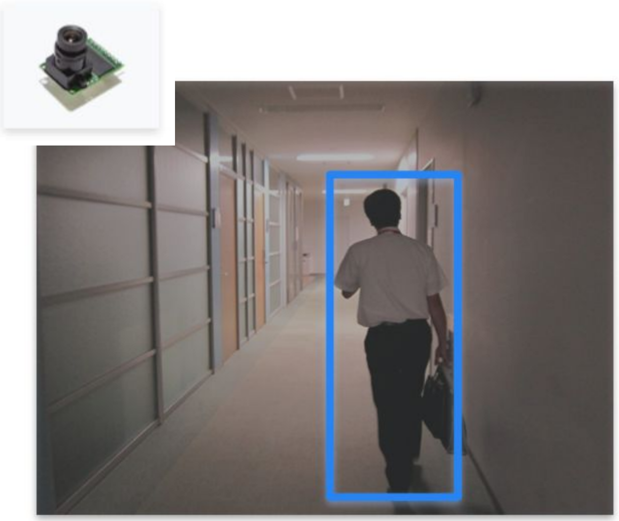
- Anomaly Detection

- Autonomous Navigation

Sound

Vibration

Vision



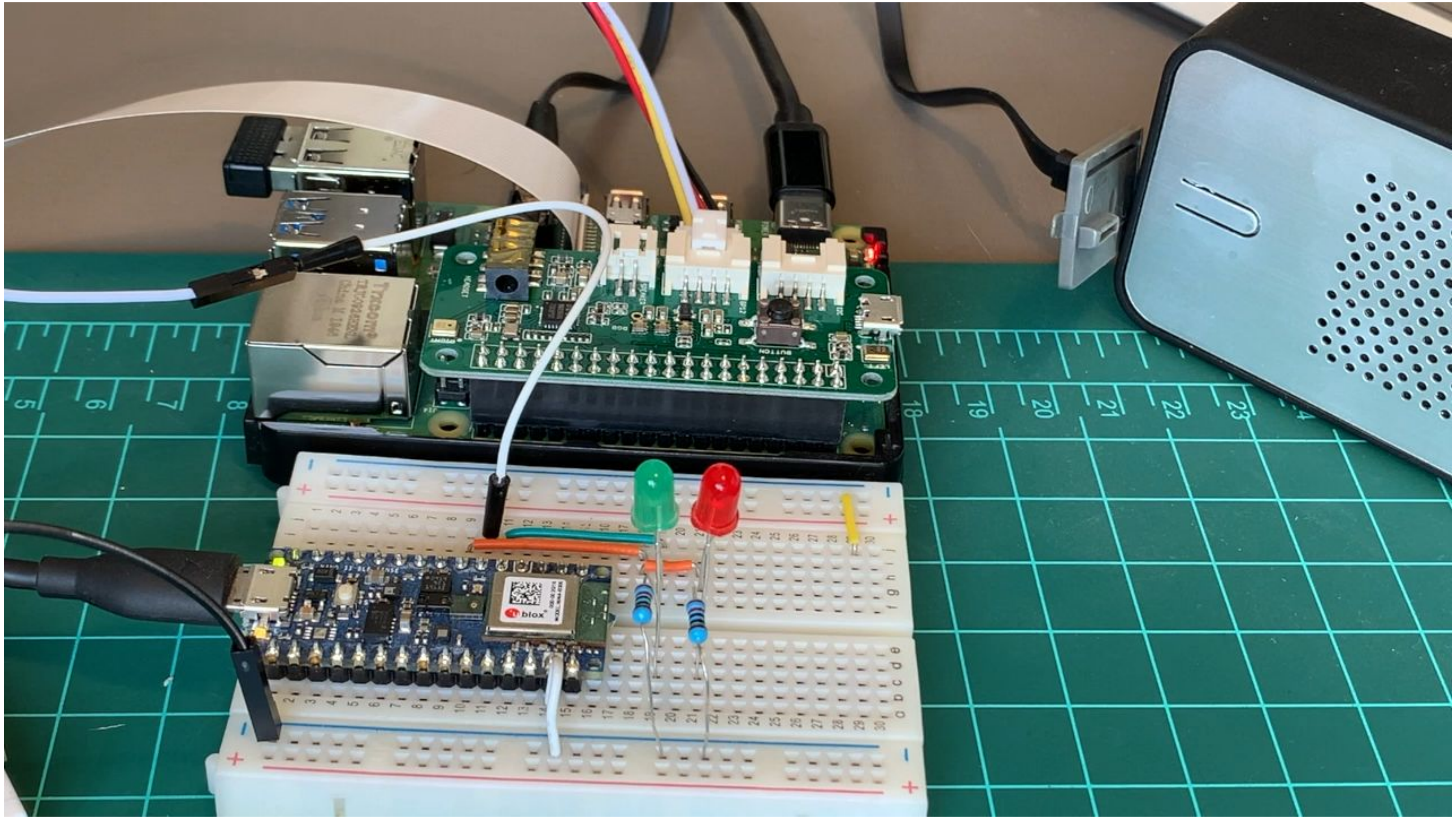
Sound

Vibration

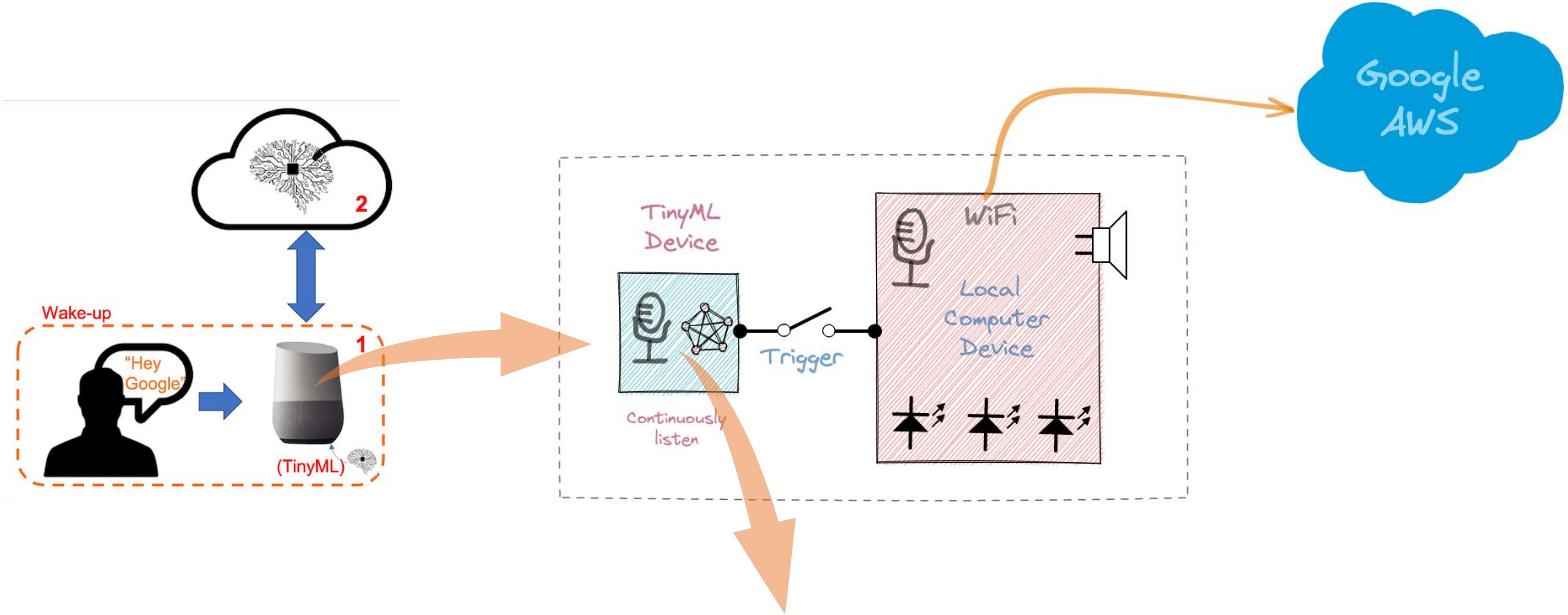
Vision

Personal Assistant

Personal Assistant

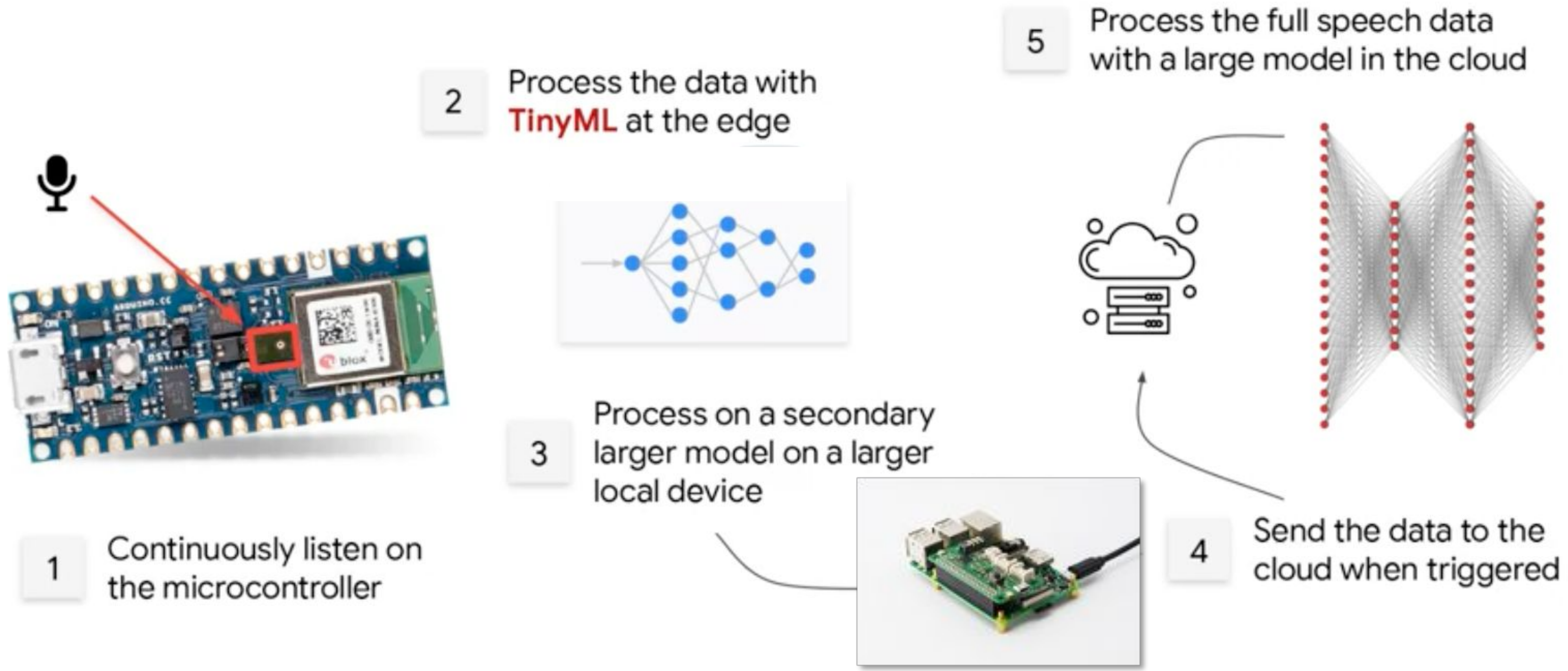


Personal Assistant

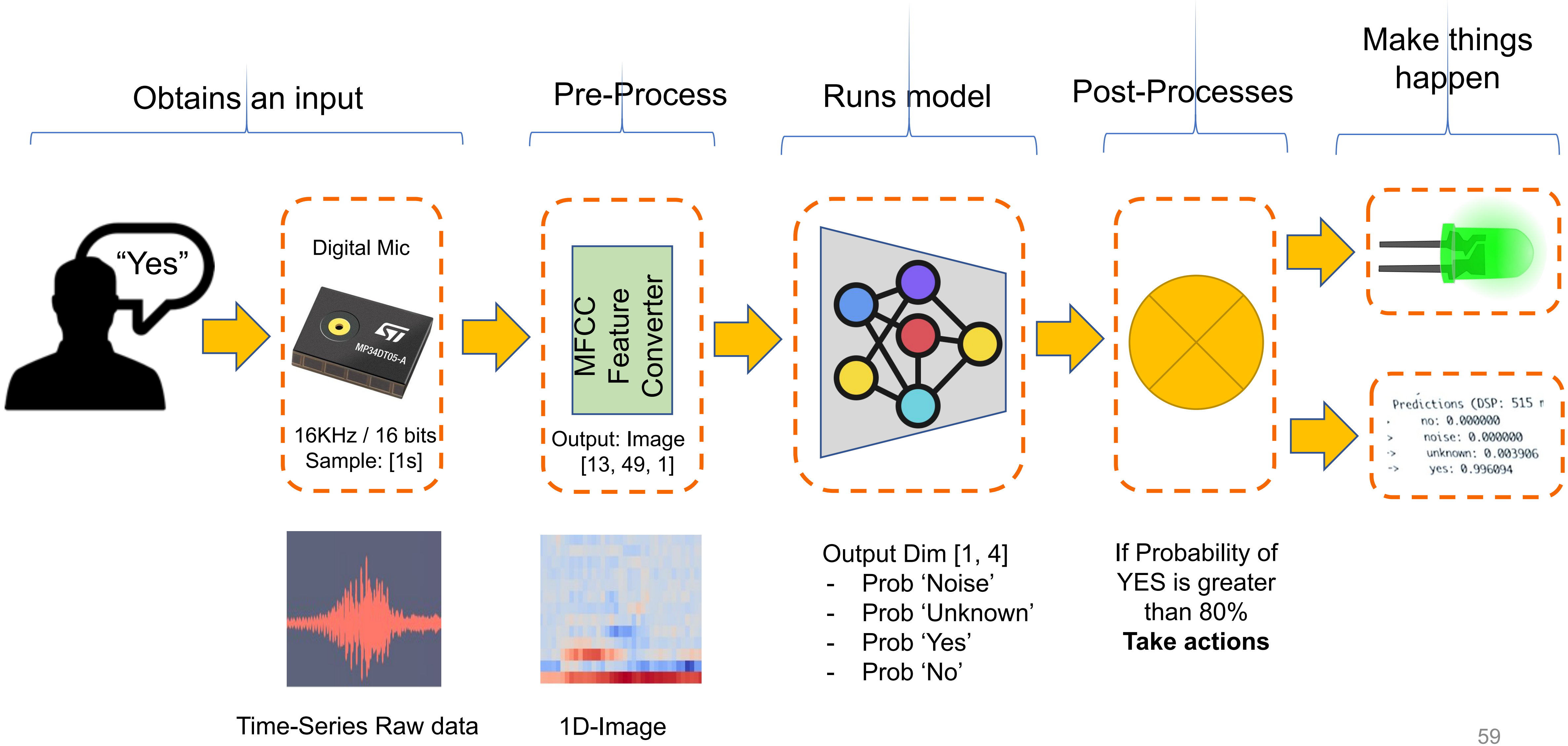


KWS

“Cascade” Detection: multi-stage model



KeyWord Spotting (KWS) - Inference



Classifying mosquito wingbeat sound using TinyML

Moez Altayeb
University of Khartoum, Sudan
ICTP, Trieste, Italy
mohedahmed@hotmail.com

Marcelo Rovai
Universidade Federal de Itajubá
Itajubá, Brazil
rovai@unifei.edu.br

Marco Zennaro
ICTP
Trieste, Italy
mzennaro@ictp.it

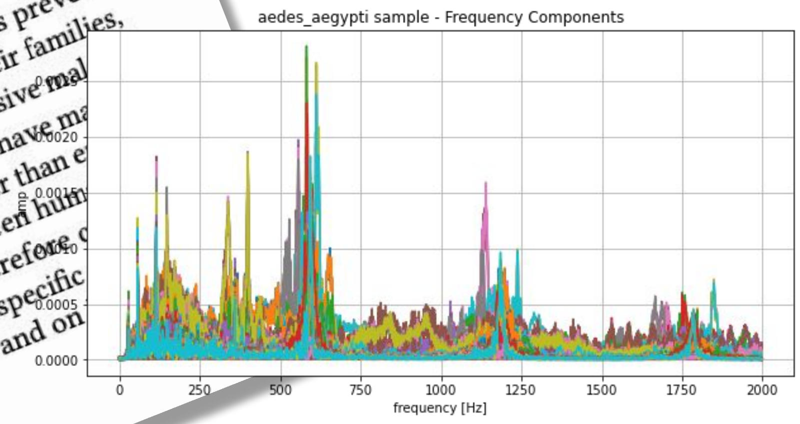
ABSTRACT

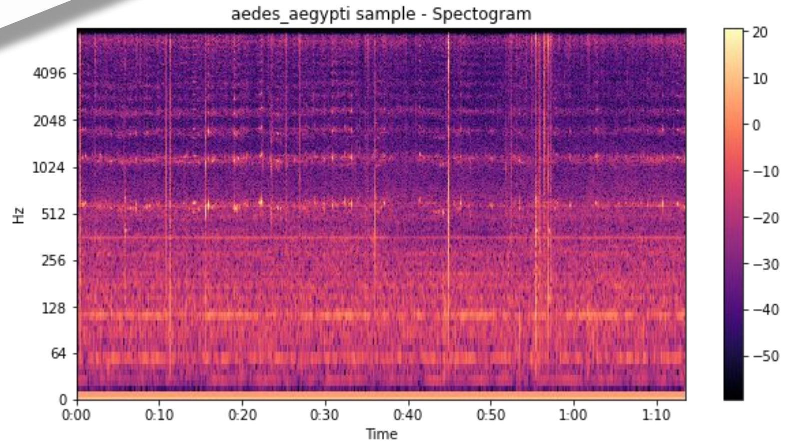
Every year more than one billion people are infected and more than one million people die from vector-borne diseases including malaria, dengue, zika and chikungunya. Mosquitoes are the best known disease vector and are geographically spread worldwide. It is important to raise awareness of mosquito proliferation by monitoring their incidence, especially in poor regions. Acoustic detection of mosquitoes has been studied for long and ML can be used to automatically identify mosquito species by their wingbeat. We present a prototype solution based on an openly available dataset, on the Edge Impulse platform and on three commercially-available TinyML devices. The proposed solution is low-power, low-cost and can run without human intervention in resource-constrained areas. This insect monitoring system can reach a global scale.

affected. People from poor communities with little access to health care and clean water sources are also at risk. Although anti-malarial drugs exist, there's currently no malaria vaccine.

Vector-borne diseases also exacerbate poverty. Illness prevent people from working and supporting themselves and their families, impeding economic development. Countries with intensive mal… have much lower income levels than those that don't have ma…

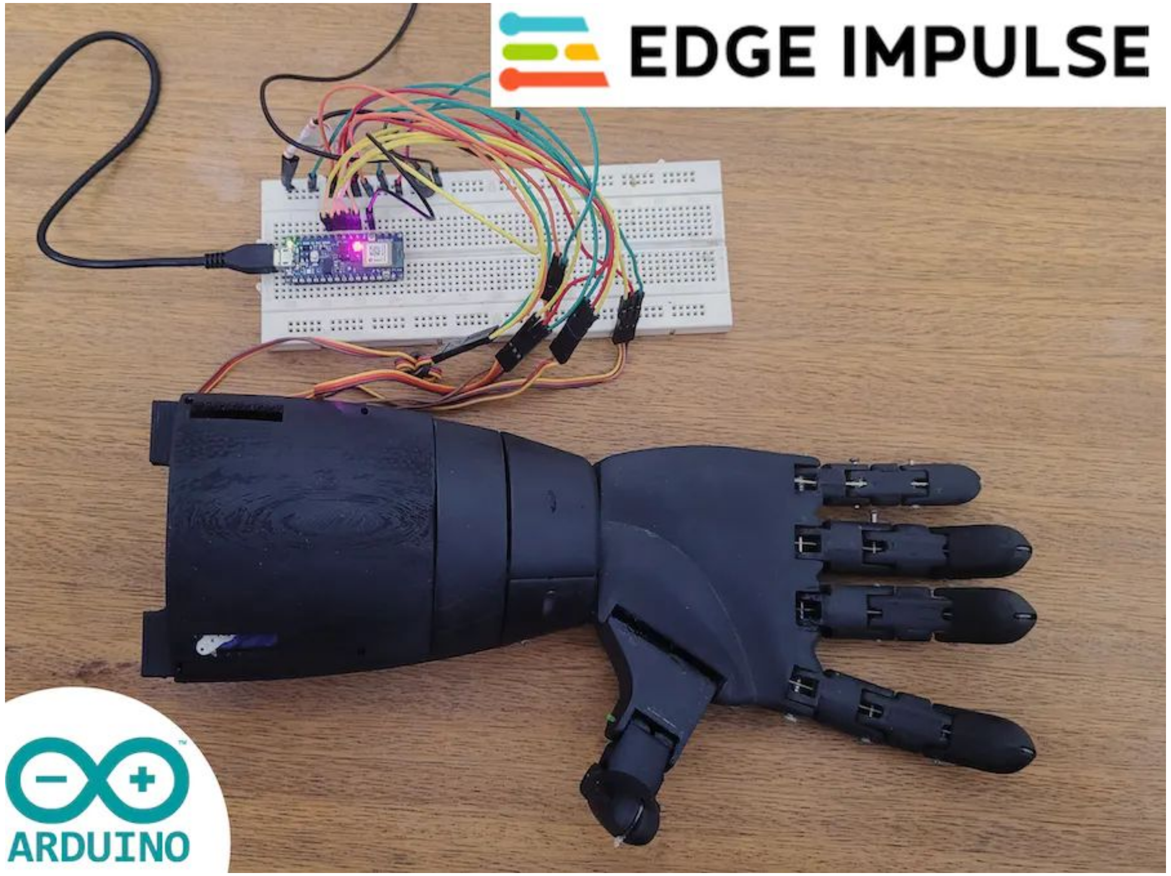
Countries affected by malaria turn to control rather than e… tion. Vector control means decreasing contact between hum… disease carriers on an area-by-area basis. It is therefore c… be able to detect the presence of mosquitoes in a specific… paper presents an approach based on TinyML and on… embedded devices.





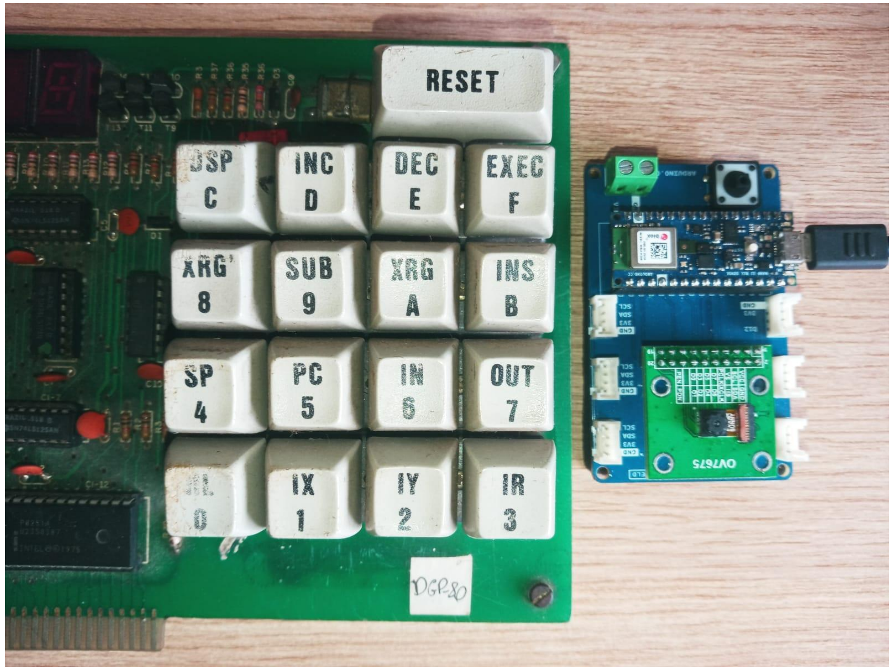
https://www.edgeimpulse.com/blog/the-world-is-all-abuzz-about-tinyml

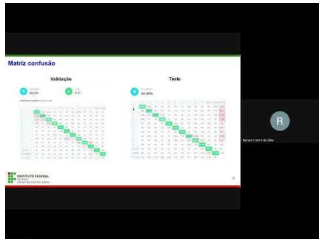
Bionic Hand Voice Commands Module



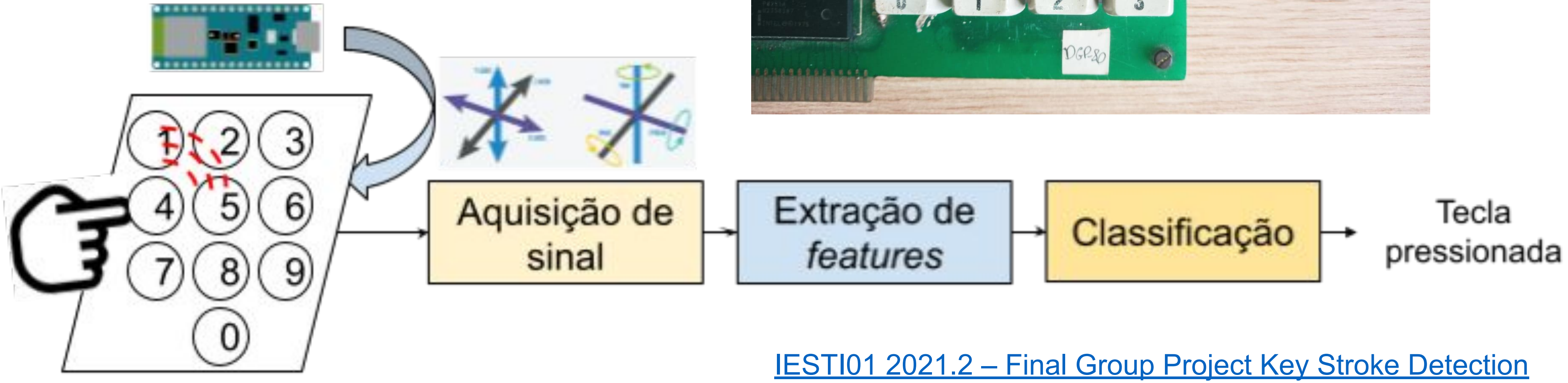
https://www.hackster.io/ex-machina/bionic-hand-voice-commands-module-w-edge-impulse-arduino-aa97e3

Keystroke **Sound** Detection





Renam Castro
Professor IFESP



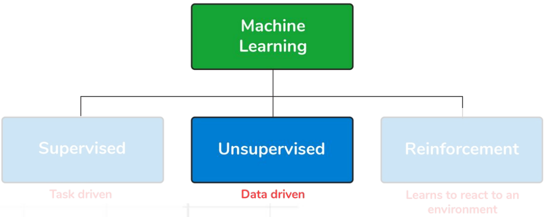
[IESTI01 2021.2 – Final Group Project Key Stroke Detection]

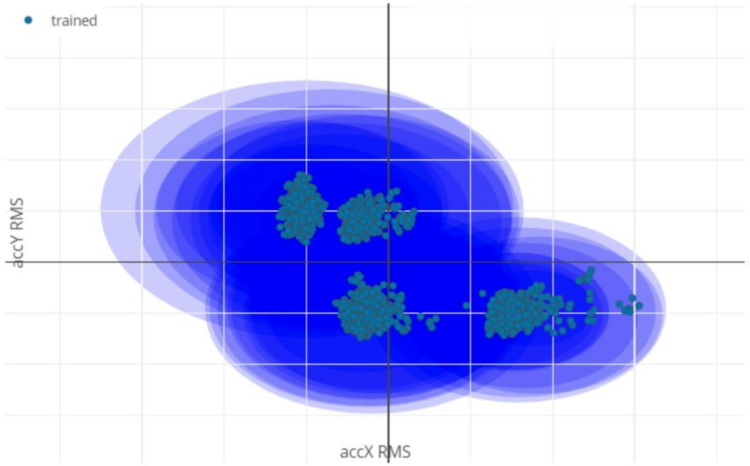
Sound

Vibration

Vision

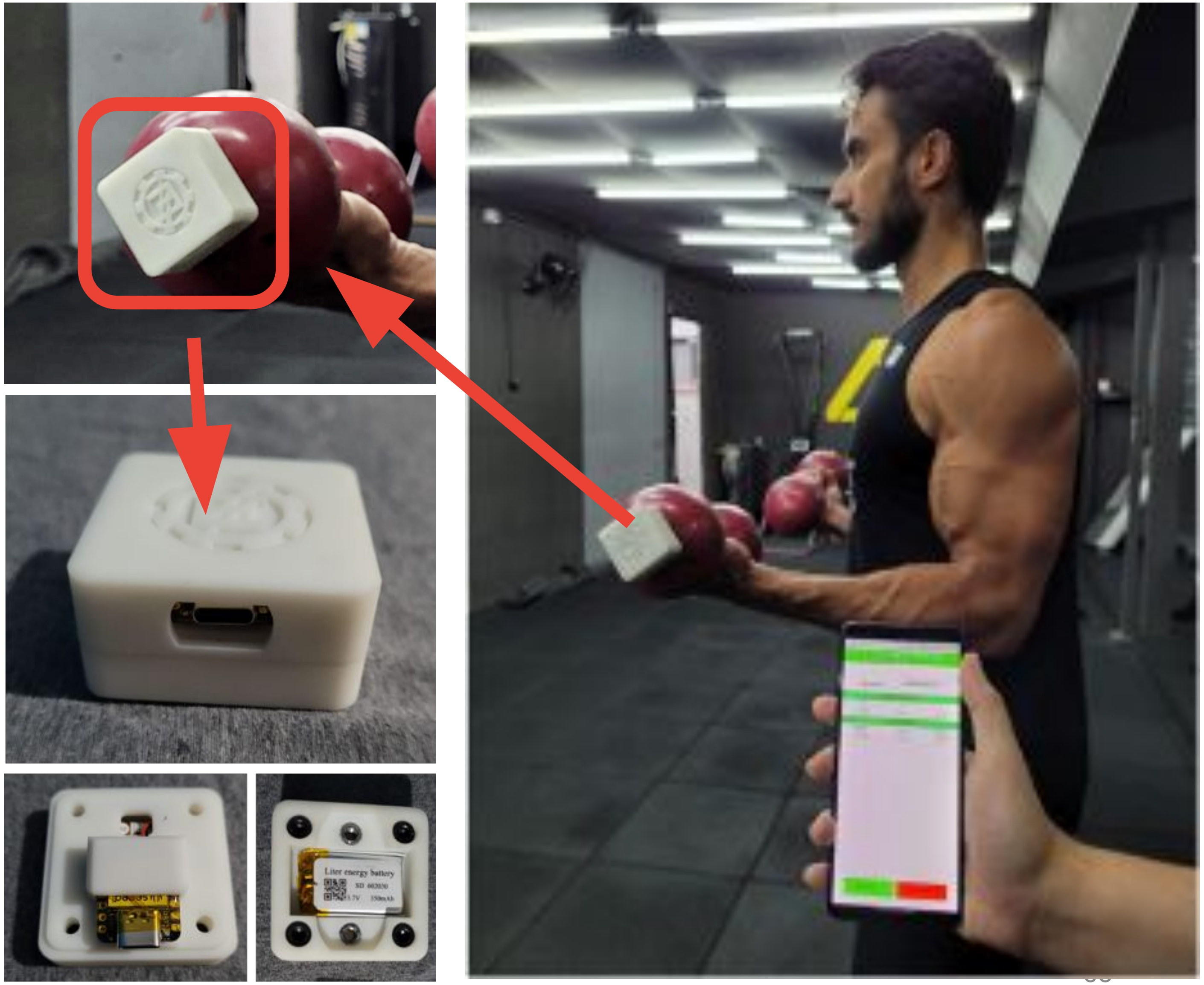
Industrial – Anomaly Detection





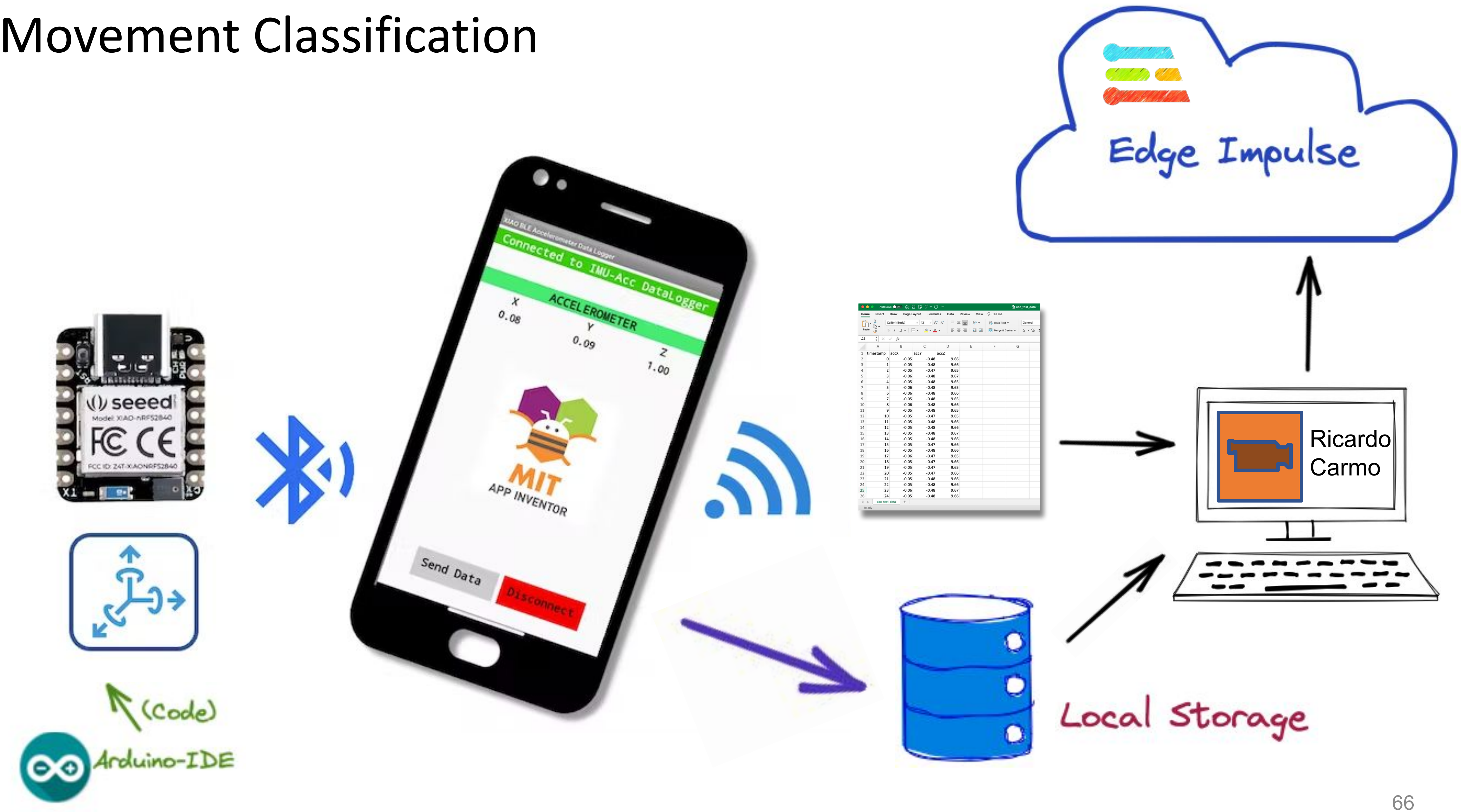
[IESTI01 2021.2 - Final Group Project: Bearing Failure Detection]

Movement Classification



CNPq UNIFEI

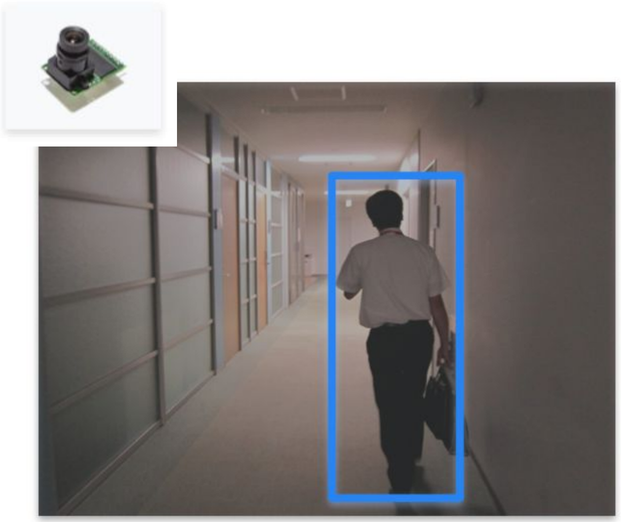
Movement Classification



Sound

Vibration

Vision



Computer Vision Main Types

Image Classification

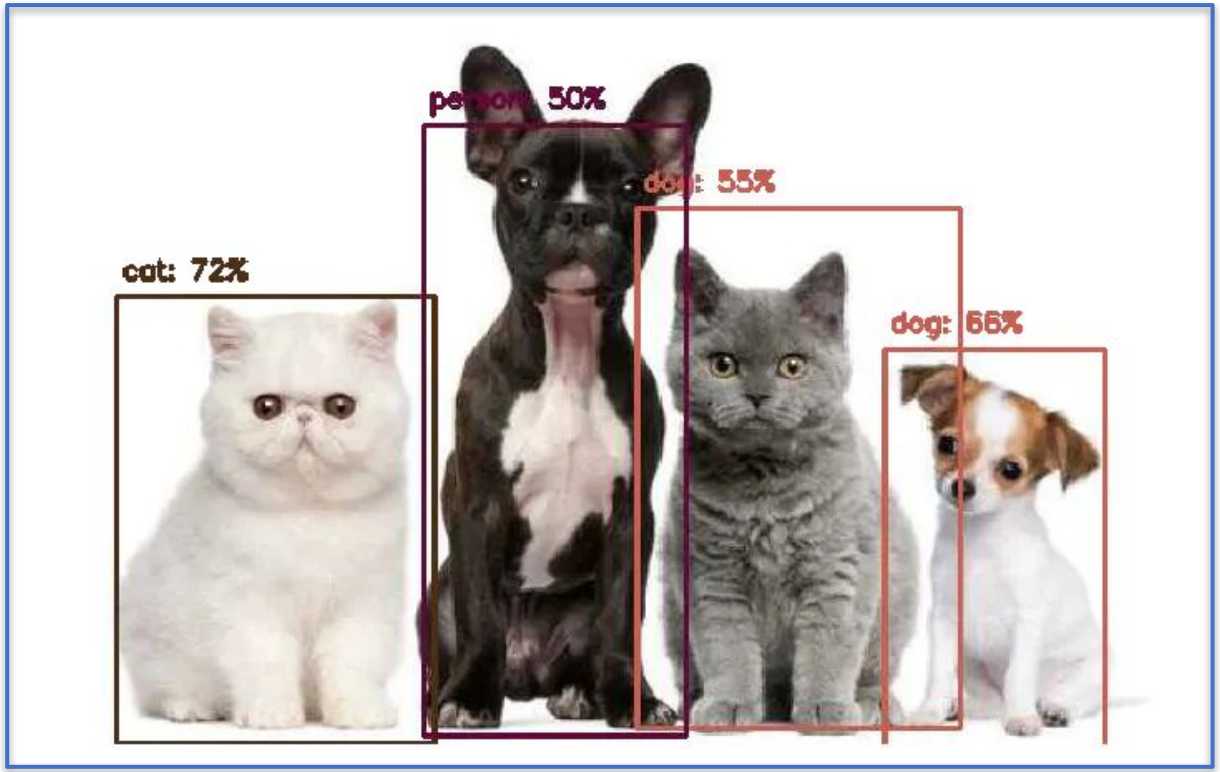
(**Multi-Class** Classification)

Cat: 70%

Dog: 80%

Object Detection

Multi-Label Classification + Object **Localization**



Computer Vision Main Types

Image Classification

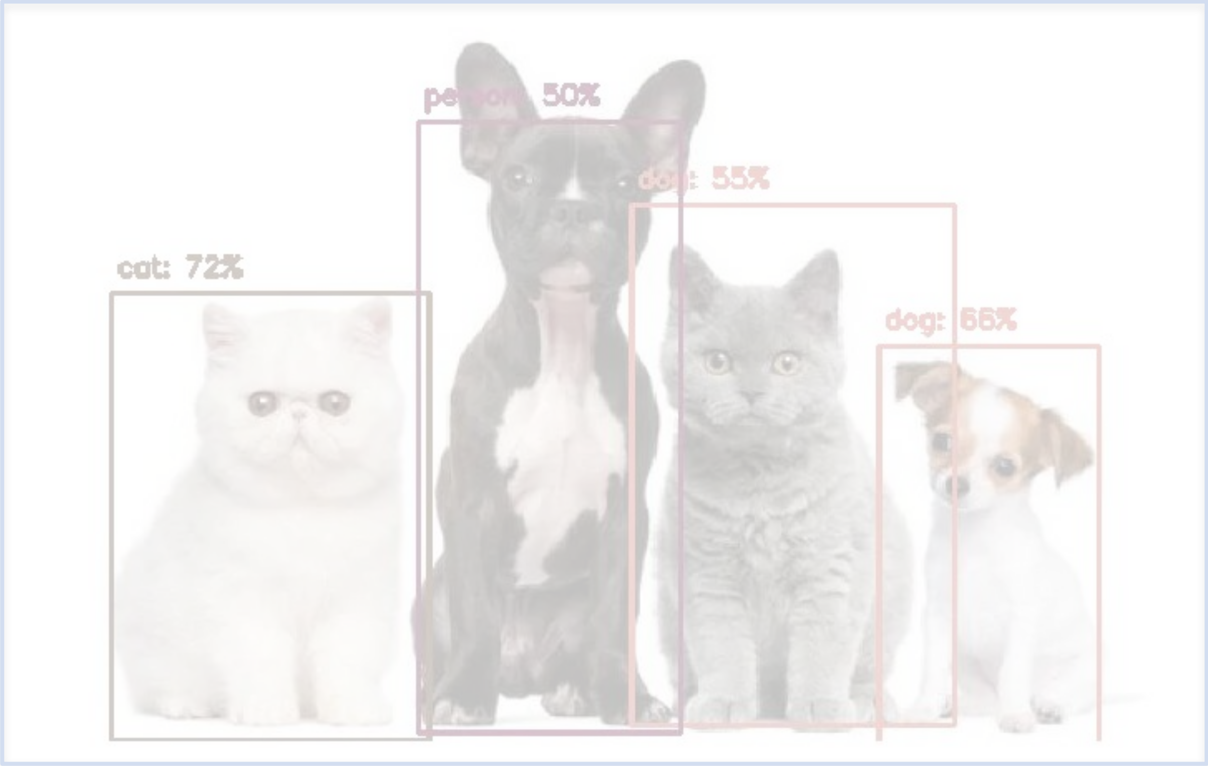
(**Multi-Class** Classification)

Cat: 70%

Dog: 80%

Object Detection

Multi-Label Classification + Object **Localization**



Forest Fire Detection

TinyML Aerial Forest Fire Detection

IESTI01 - Forest Fire Detection – Proof of Concept

Coffee Disease Classification

João Vitor Yukio Bordin Yamashita
Graduando em Engenharia Eletrônica pela UNIFEI

https://www.hackster.io/Yukio/coffee-disease-classification-with-ml-b0a3fc

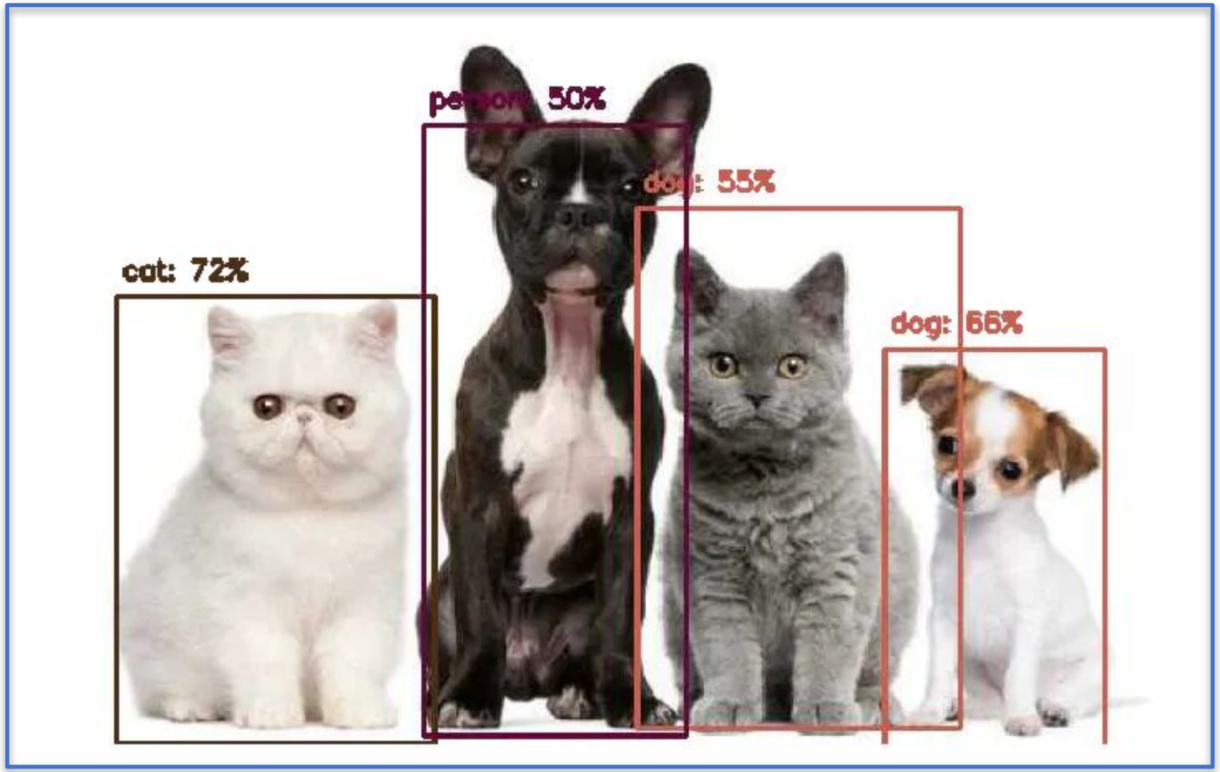
Computer Vision Main Types

Image Classification
(**Multi-Class** Classification)

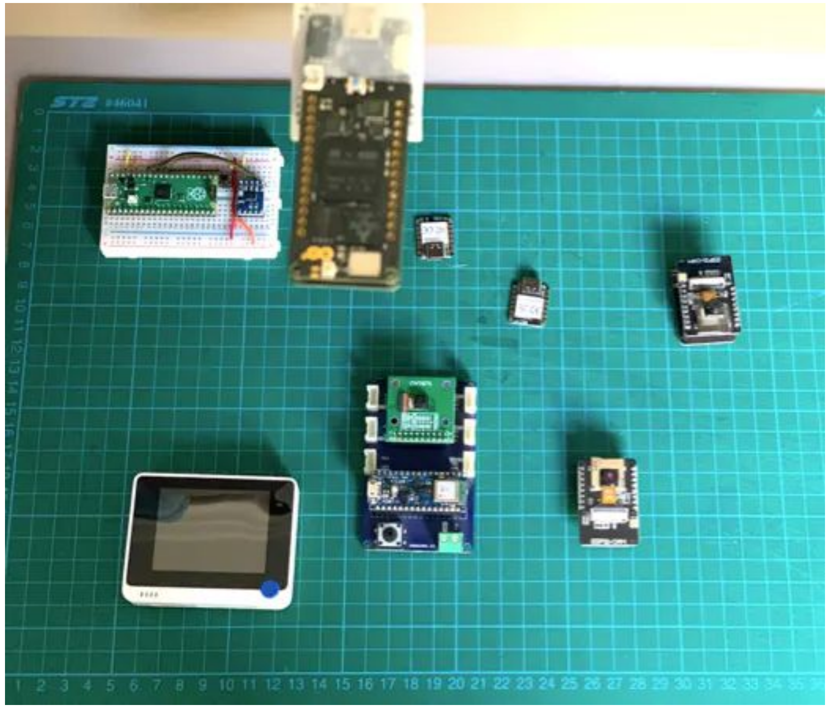
Cat: 70%

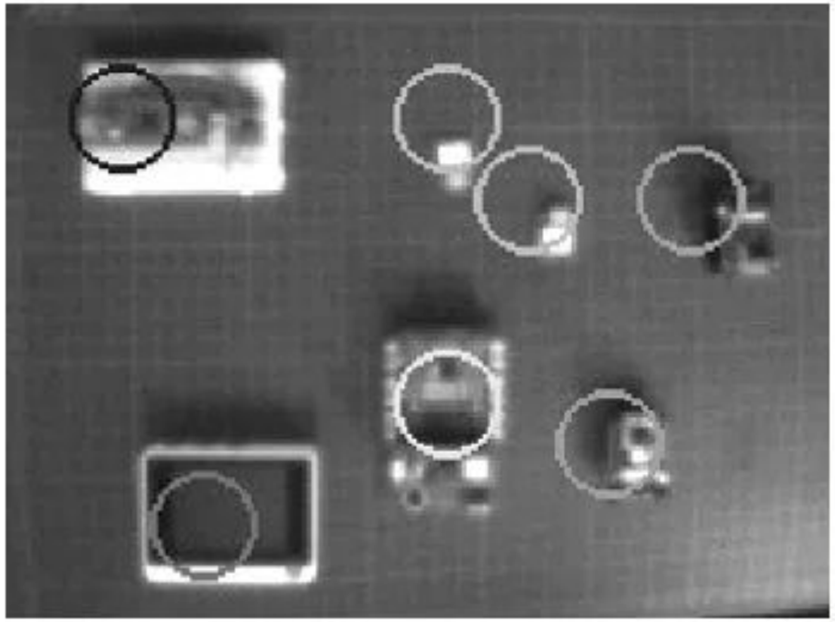
Dog: 80%

Object Detection
Multi-Label Classification + Object **Localization**



Detecting Objects using TinyML (FOMO)

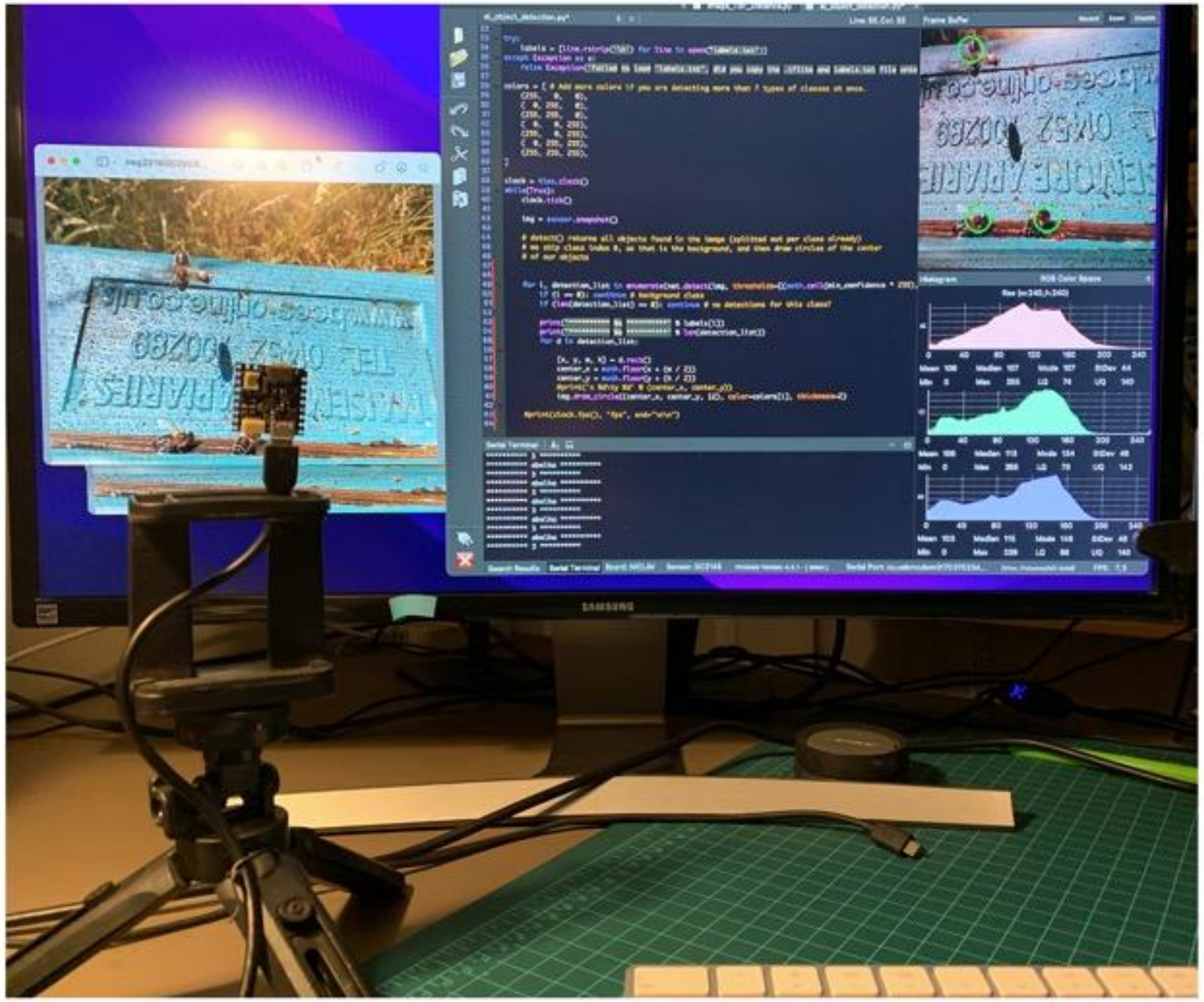




```
********** espcam **********
x 70    y 150
x 130   y 170
********** nano **********
x 70    y 110
********** pico **********
x 150   y 30
********** wio **********
x 50    y 50
********** xiao **********
x 150   y 110
x 130   y 130
6.97512 fps
```

[EdgeAI made simple - Exploring Image Processing (Object Detection) on microcontrollers with Arduino Portenta, Edge Impulse FOMO, and OpenMV]

Detecting Objects using TinyML (FOMO)

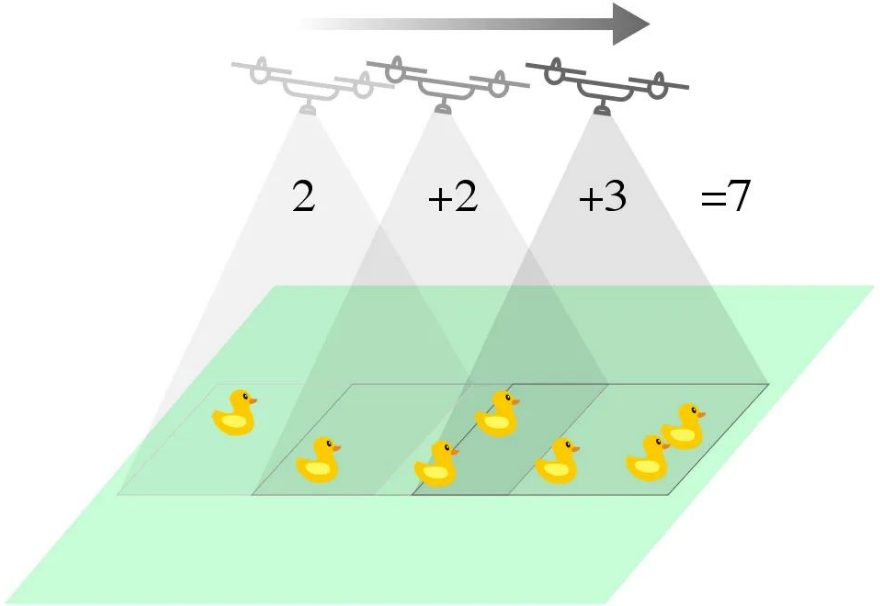


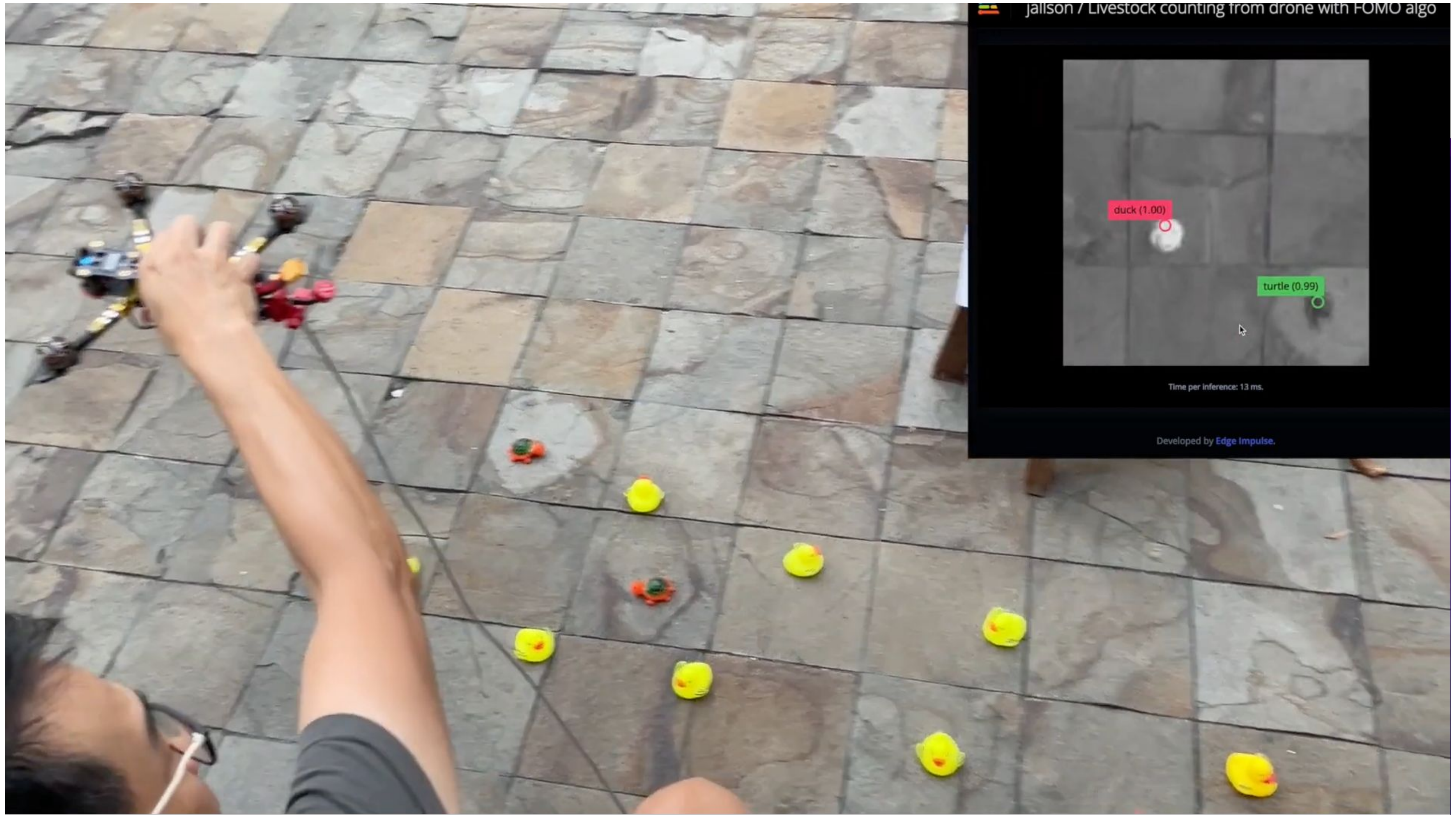
OpenMV

MicroPython

EDGE IMPULSE

Livestock / Wildlife Counting from Drone with FOMO





https://www.hackster.io/jallsonsuryo/livestock-wildlife-counting-from-drone-with-fomo-algorithm-a2f734

Other TinyML / MCUs Project Examples

Vision

- Image Classification with ESP32-CAM [Doc]
- Image Classification with Portenta H7 [Doc]
- Object Detection with Portenta H7 [Doc]

Sound

- Listening Temperature with Nano 33 [Doc]
- COPD Detection with Nano 33 [Doc]
- Sound Classification with XIAO BLE Sense [Doc]

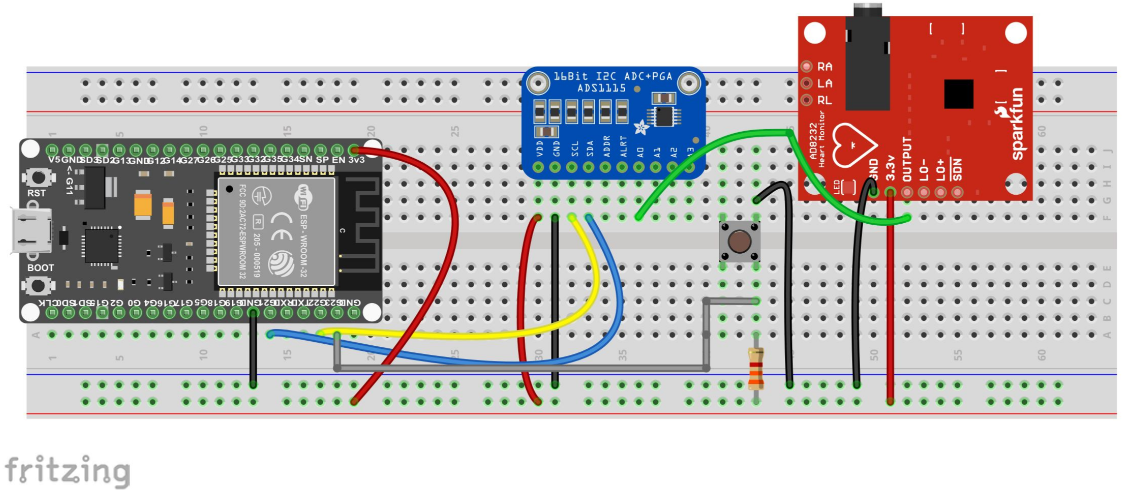
Vibration

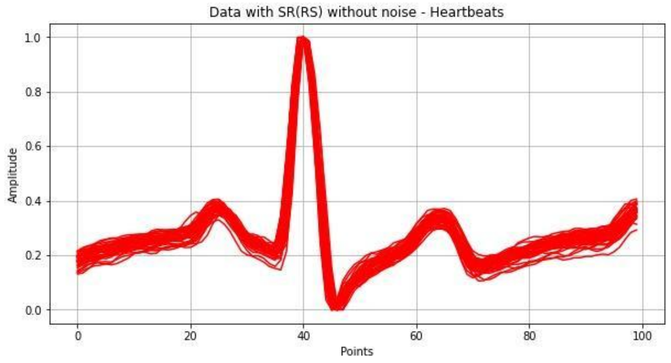
- Motion Recognition with RPi Pico [Doc]
- Gesture Recognition with Wio Terminal [Doc]
- Anomaly Detection with XIAO BLE Sense [Doc]

Other Sensors / MCUs / Models

Examples

AD8232 - Single Lead Heart Rate Monitor

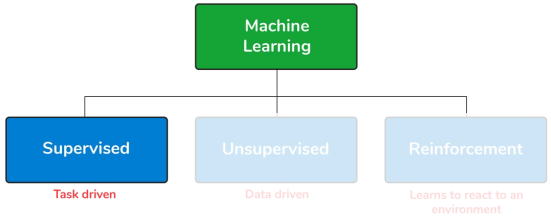


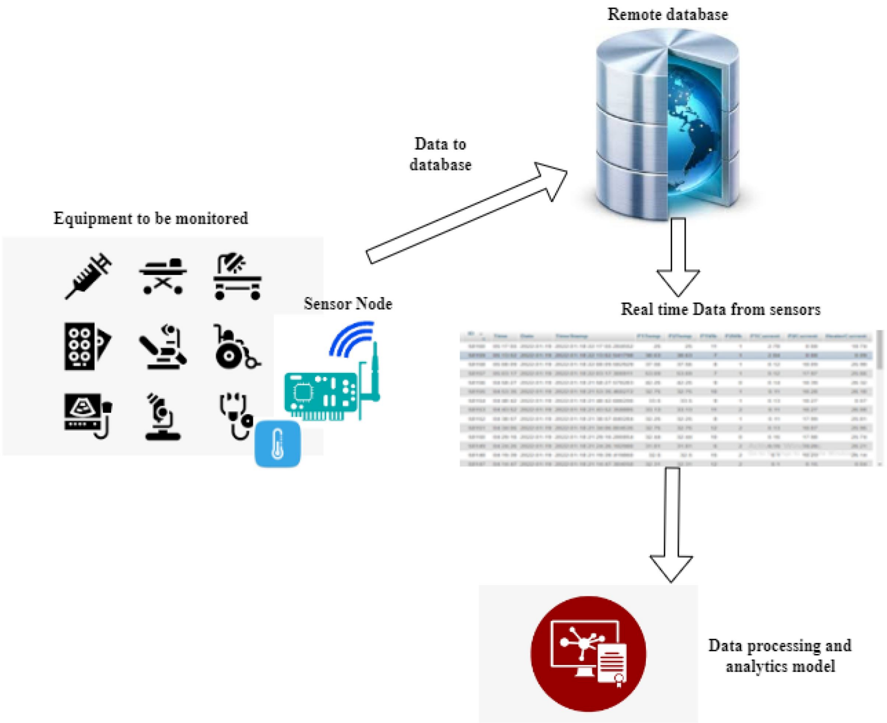


Guilherme Silva
Engenheiro - UNIFEI

[Atrial Fibrillation Detection on ECG using TinyML
Silva et al. UNIFEI 2021]

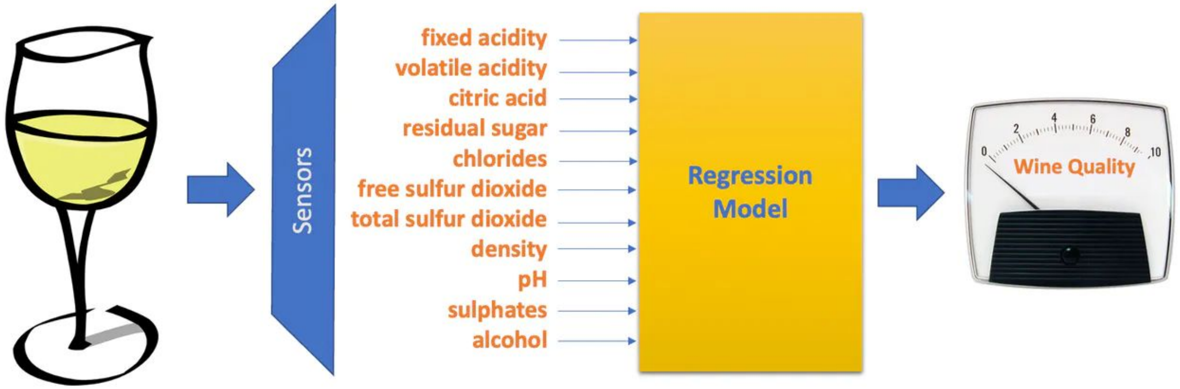
Regression on TinyML





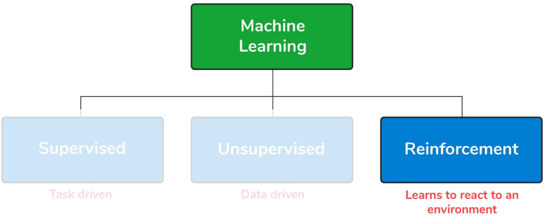
[On-Device IoT-Based Predictive Maintenance Analytics Model: Comparing TinyLSTM and TinyModel from Edge Impulse]()

Sensor fusion



[TinyML Made Easy: Exploring Regression - White Wine Quality]()

Reinforcement on TinyML



Deep Reinforcement Learning for Autonomous Source Seeking on a Nano Drone

Bardienus P. Duisterhof[1,3] Srivatsan Krishnan[1] Jonathan J. Cruz[1] Colby R. Banbury[1] William Fu[1]

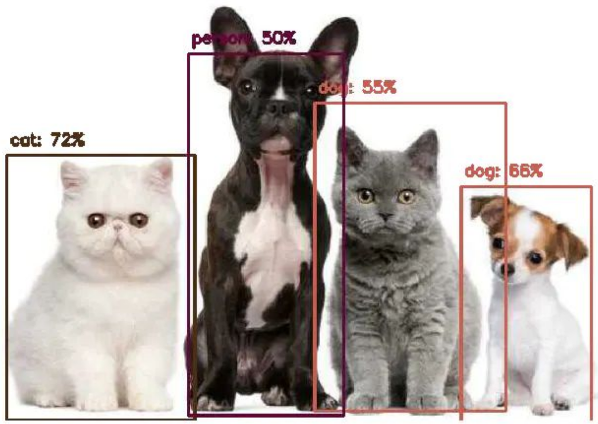
Aleksandra Faust[2] Guido C. H. E. de Croon[3] Vijay Janapa Reddi[1,4]

[1]Harvard University, [2]Robotics at Google, [3]Delft University of Technology, [4]The University of Texas at Austin

https://arxiv.org/abs/1909.11236

https://youtu.be/wmVKbX7MOnU

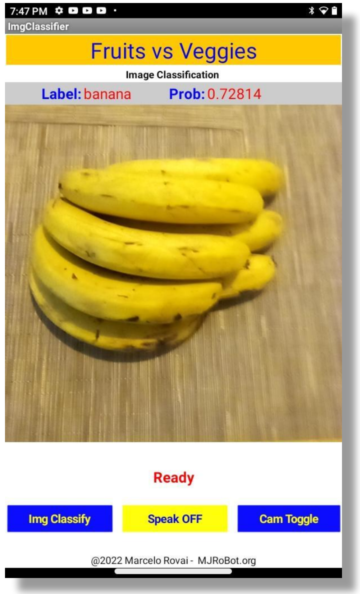
Exploring AI at the edge (Computer Vision)

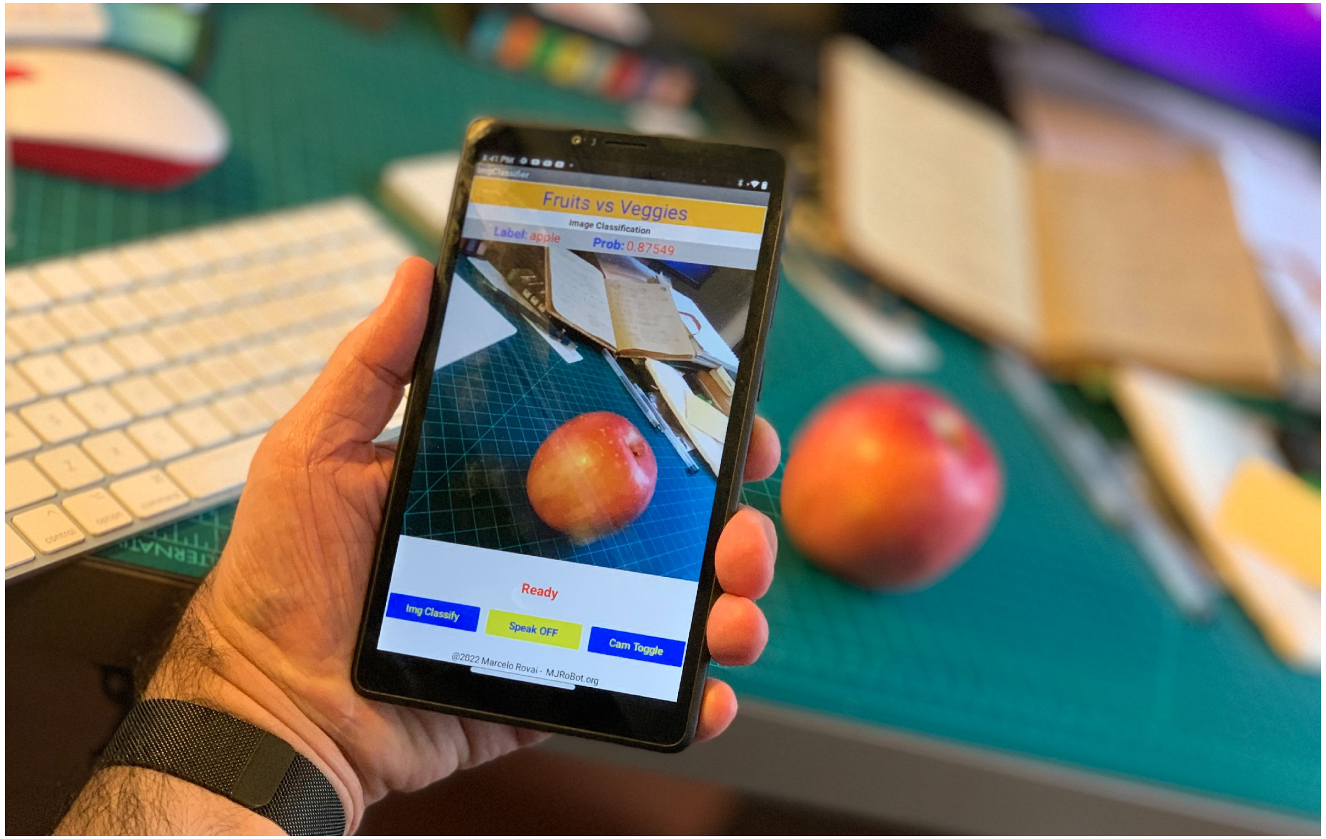


[Exploring AI at the Edge!]()

Classifying Images using Smartphones

AI WITH APP INVENTOR





https://www.hackster.io/mjrobot/app-inventor-edgeml-image-classification-fruit-vs-veggies-b671da

Thanks

TINYML4D