IoT 2.0 – When IoT meets the AI

Bringing intelligence to sensors

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

UNIFEI - Universidade Federal de Itajubá, Brazil

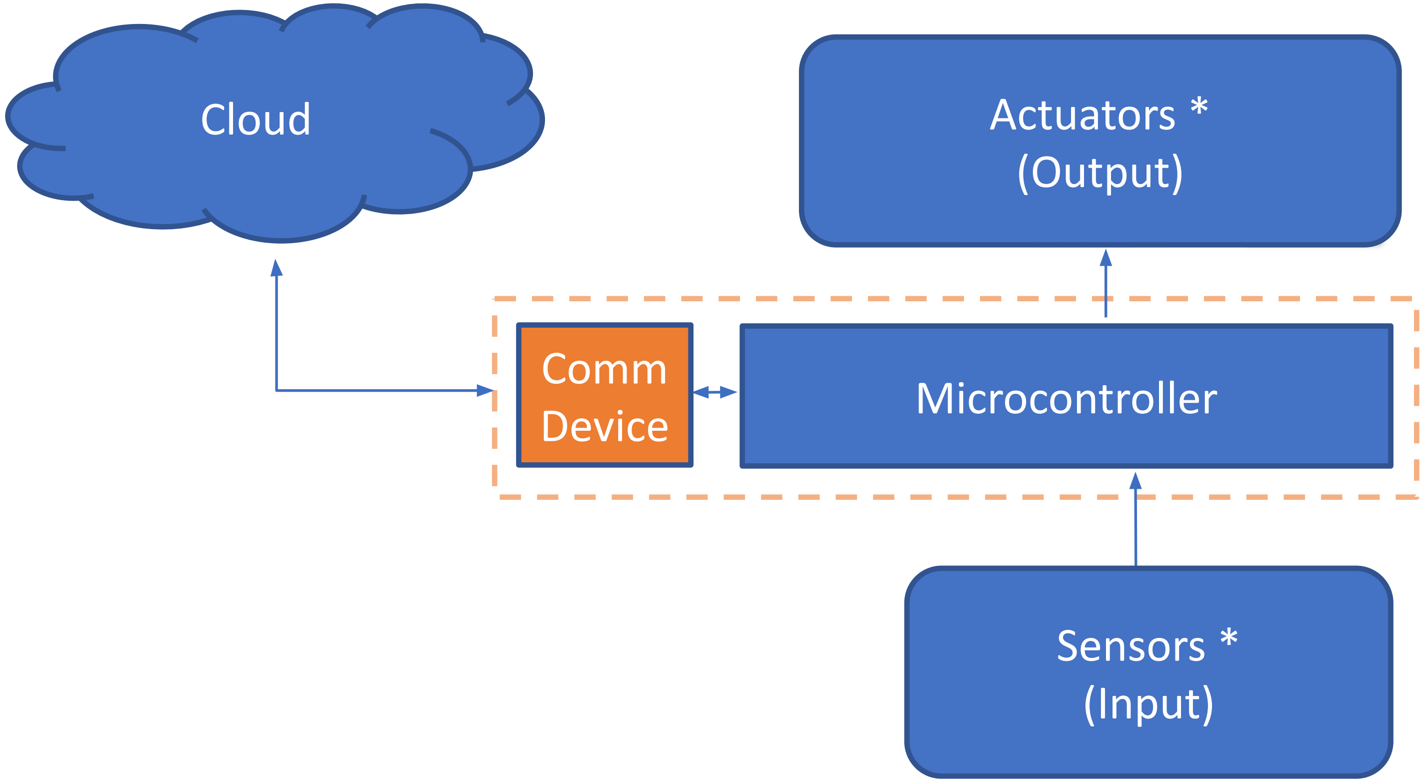
Marcelo Rovai was born in São Paulo and held a Master's degree in Data Science from the Universidad del Desarrollo (UDD) in Chile and an MBA from IBMEC (INSPER) in Brazil. He graduated in 1982 as an Engineer from UNIFEI, Federal University of Itajuba, with a specialization from Escola Politécnica de Engenharia of São Paulo University (USP); both institutions are located in Brazil.

Mr. Rovai has experience as a teacher, engineer, and executive in several technology companies such as CDT/ETEP, AVIBRAS Aeroespacial, SID Informática, ATT-GIS, NCR, DELL, COMPAQ (HP), and more recently at IGT as a VP. He now works at IGT as a Senior Advisor for Latin America.

Marcelo Rovai publishes articles about electronics on websites such as MJRoBot.org, Hackster.io, Instructables.com, and Medium.com. Furthermore, he is a volunteer Professor at the UNIFEI Engineering Institute in Brazil and a lecturer at several Congresses and Universities on the topics of IoT and TinyML. He is an active member and a Co-Chair of the TinyML4D group, an initiative to bring TinyML education to developing countries.

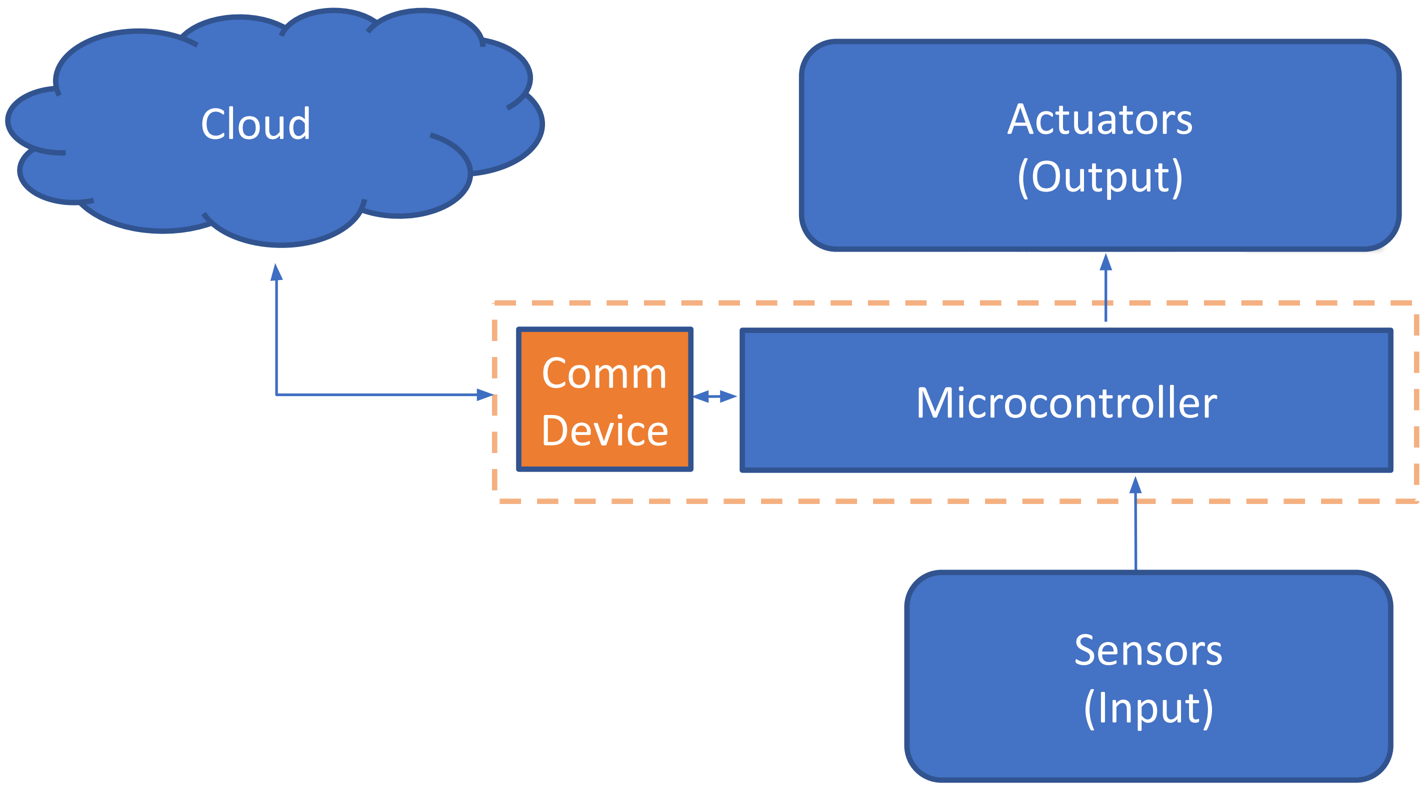
Internet of Things (IoT)

Typical IoT Project



* “Things”

Typical IoT Project



5 Quintillion

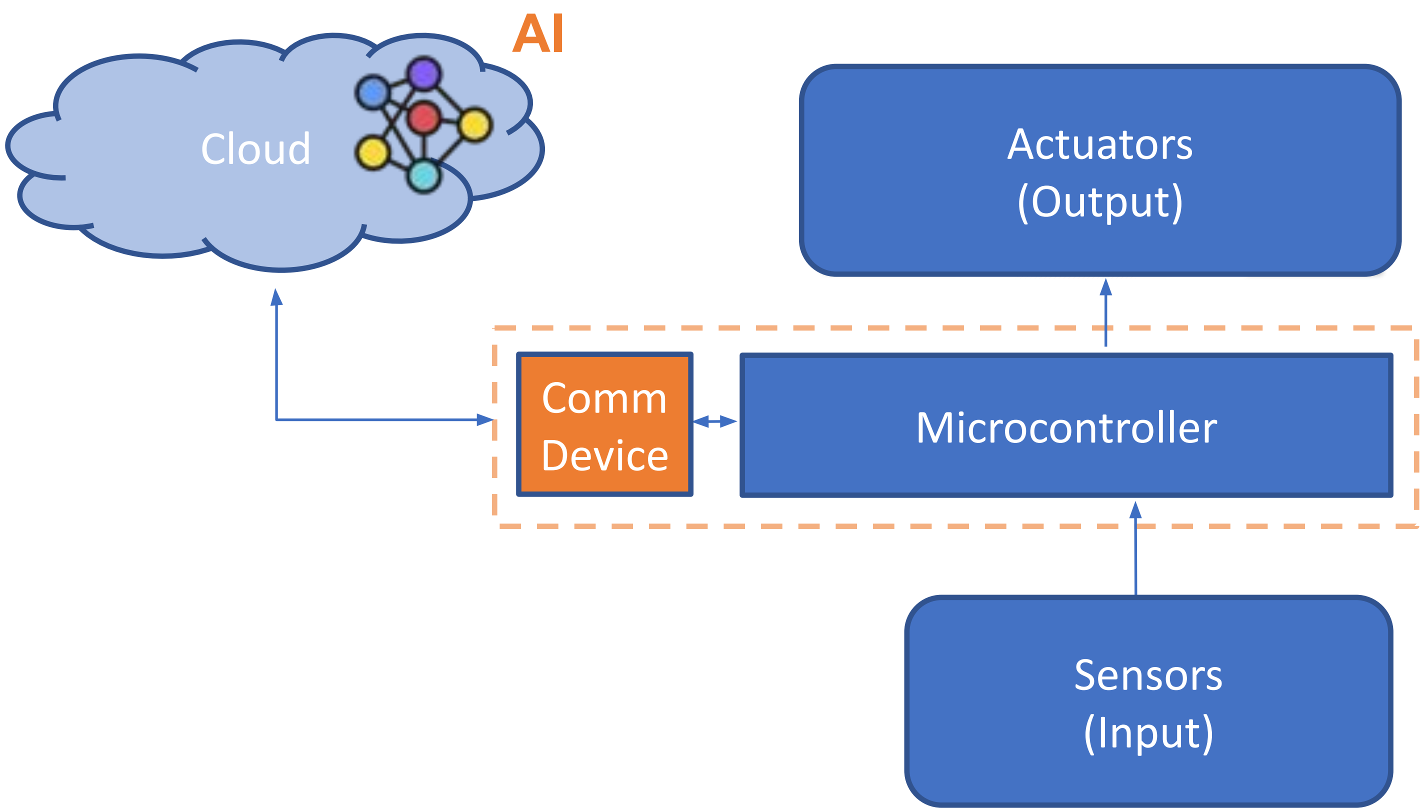
bytes of data produced every day by IoT

<1%

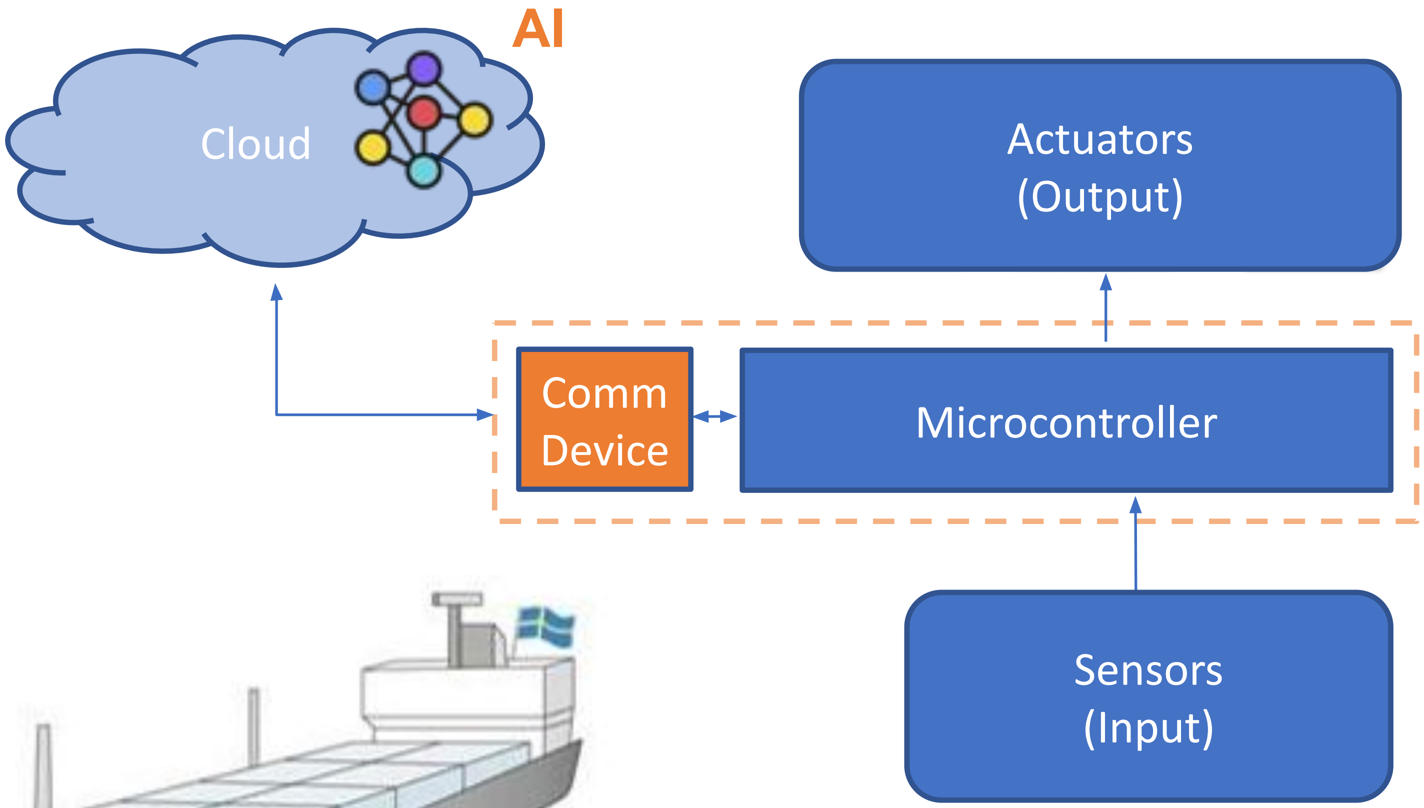
of unstructured data is analyzed or used at all

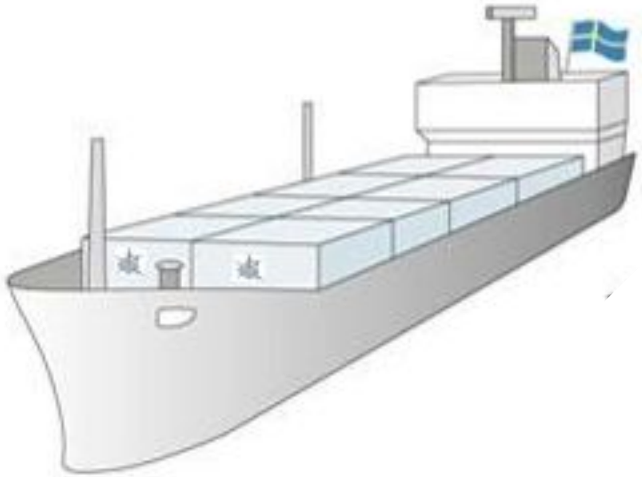
Source: Harvard Business Review, [What's Your Data Strategy?](), April 18, 2017
Cisco, [Internet of Things (IoT) Data Continues to Explode Exponentially. Who Is Using That Data and How?](), Feb 5, 2018

Typical AIoT Project



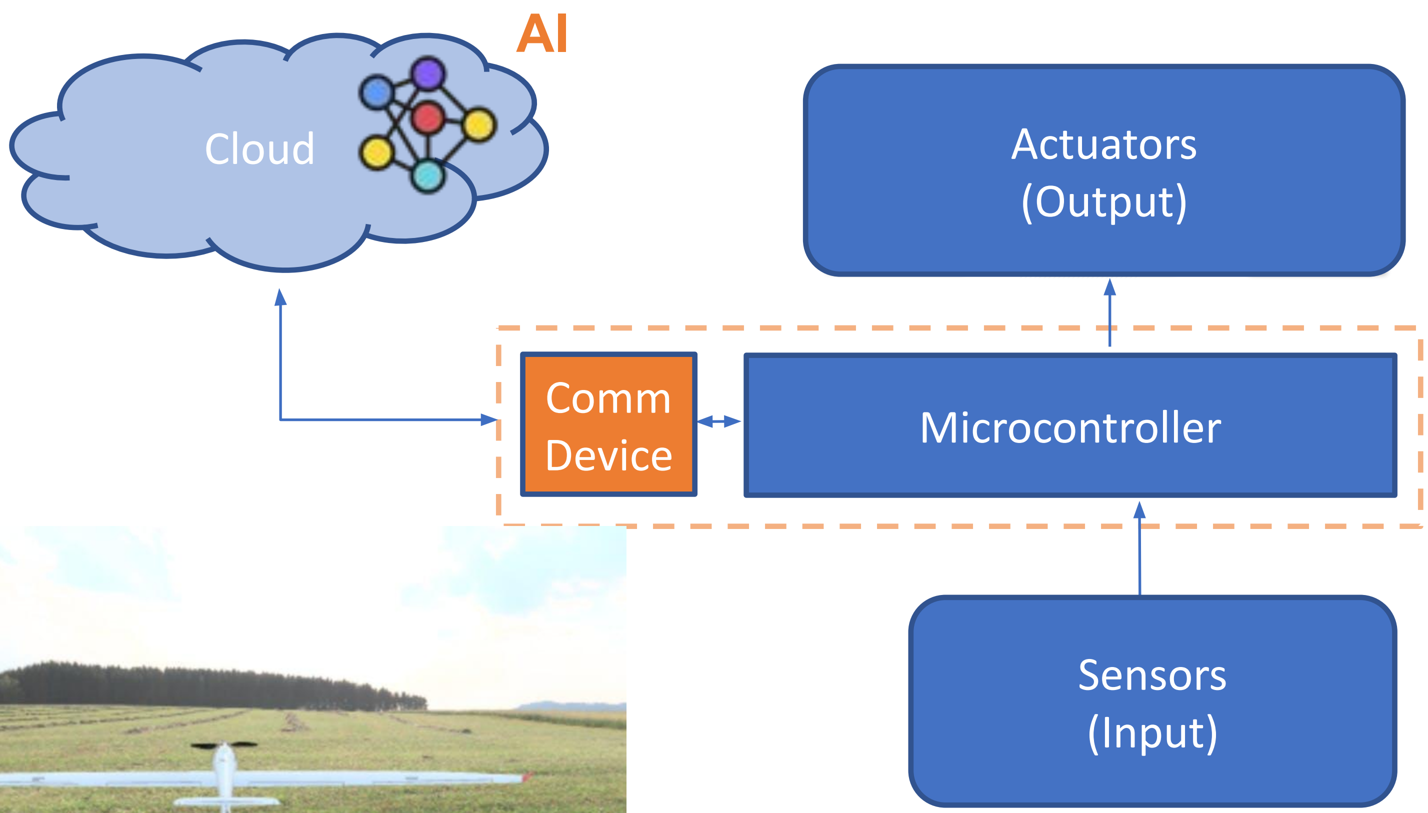
Typical AIoT Project … … Issues





Bandwidth

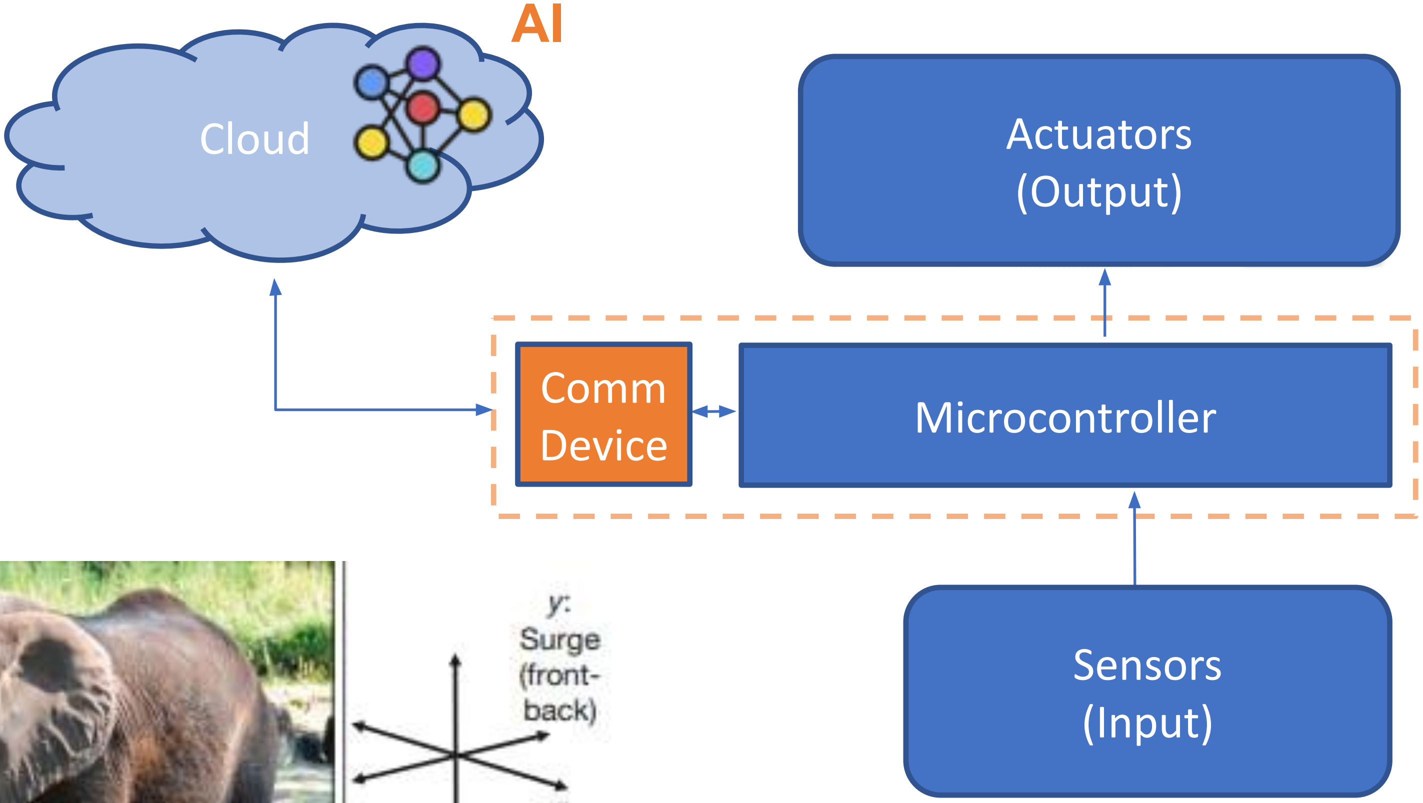
Typical AIoT Project … … Issues

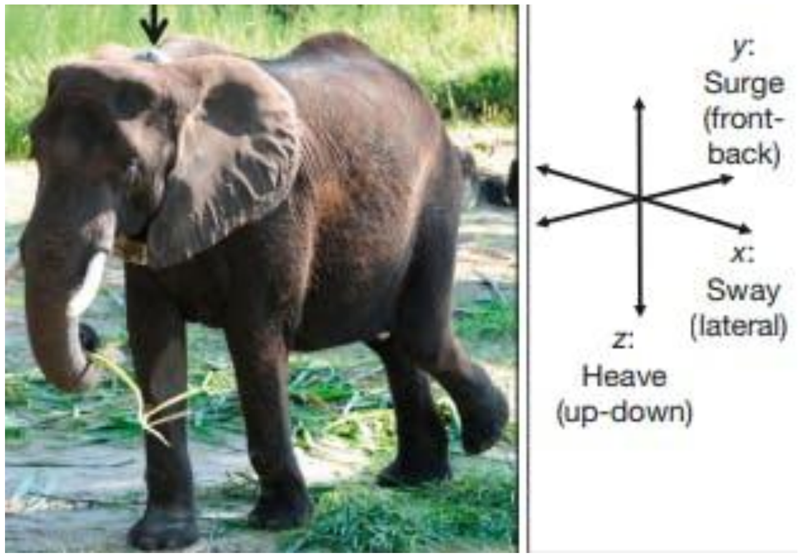


Bandwidth

Latency

Typical AIoT Project … … Issues



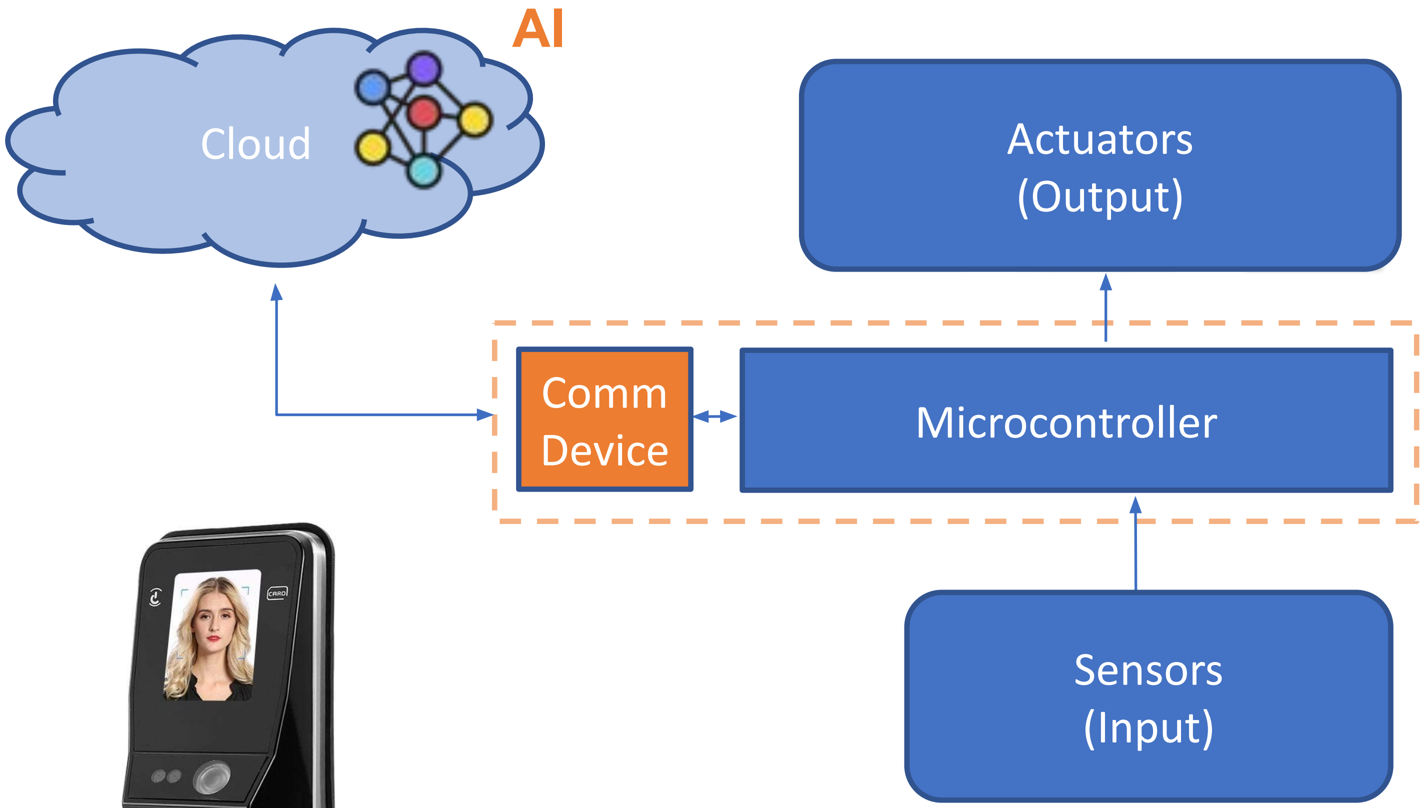


Bandwidth

Latency

Energy

Typical AIoT Project … … Issues



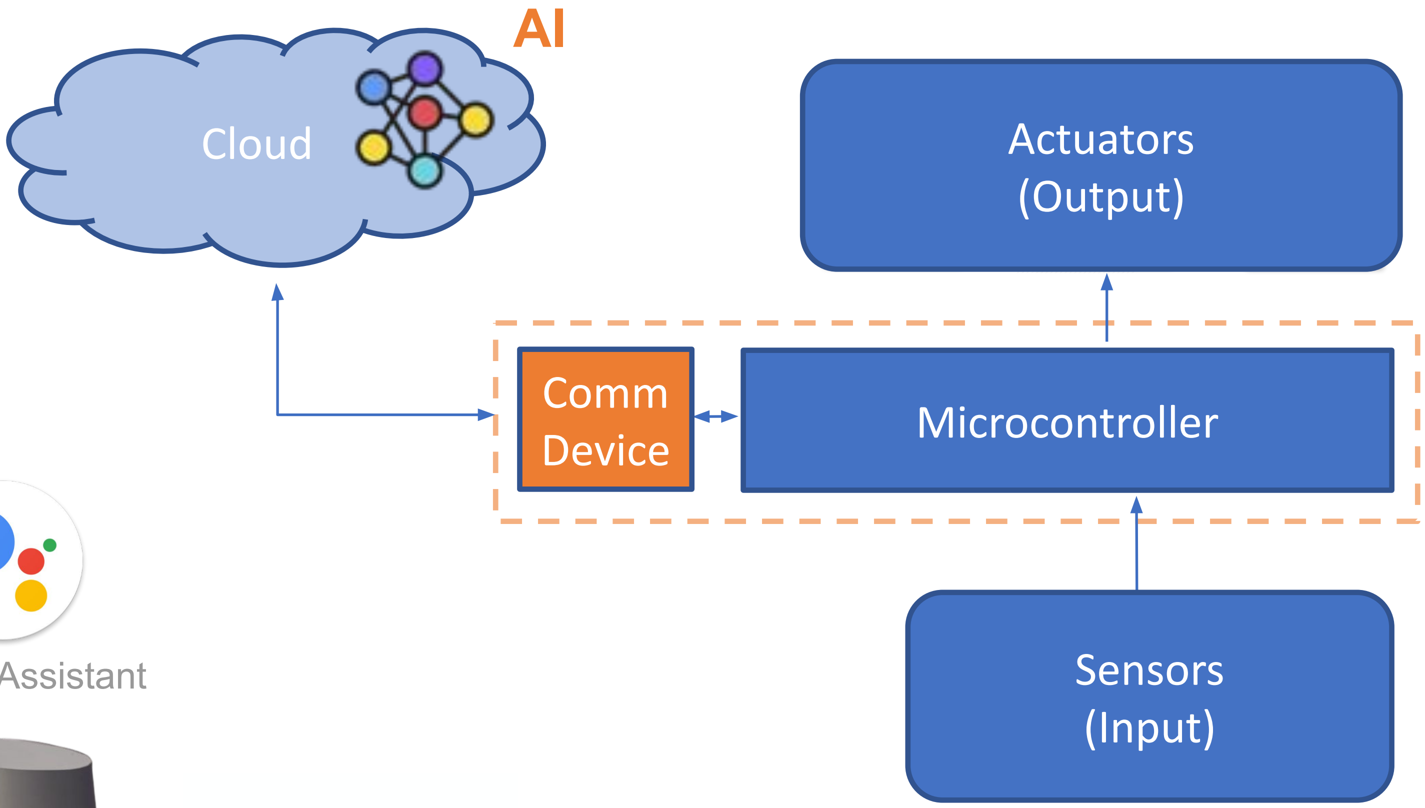
Bandwidth

Latency

Energy

Reliability

Typical AIoT Project … … Issues



Google Assistant

Bandwidth

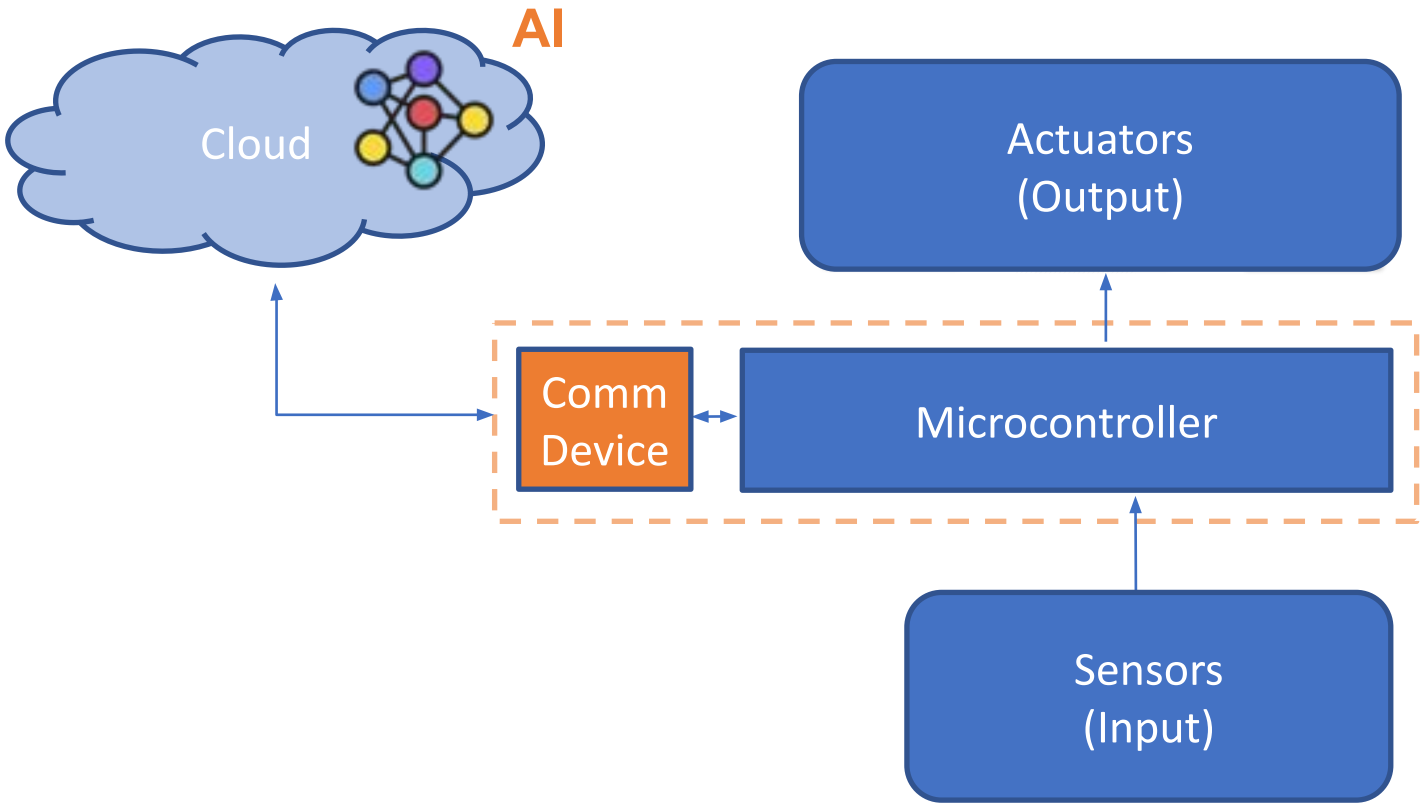
Latency

Energy

Reliability

Privacy

Typical AIoT Project … … Issues

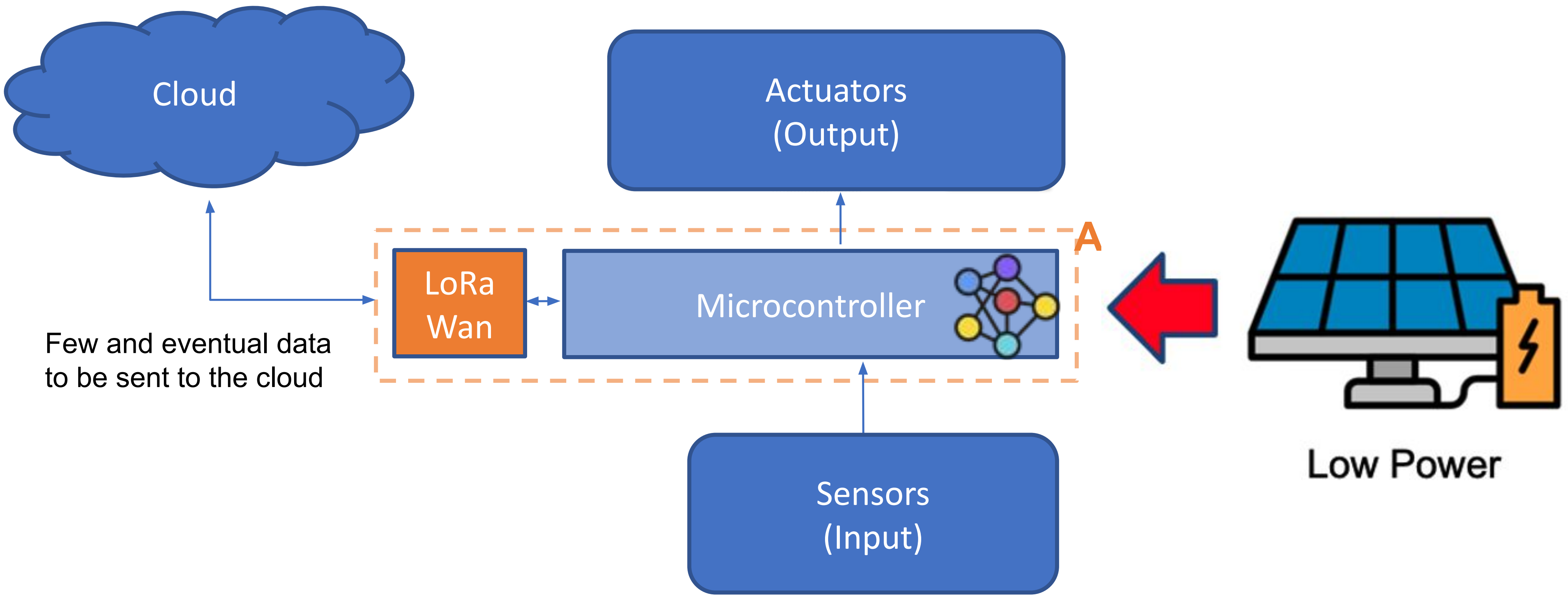


Bandwidth
Latency
Energy
Reliability
Privacy

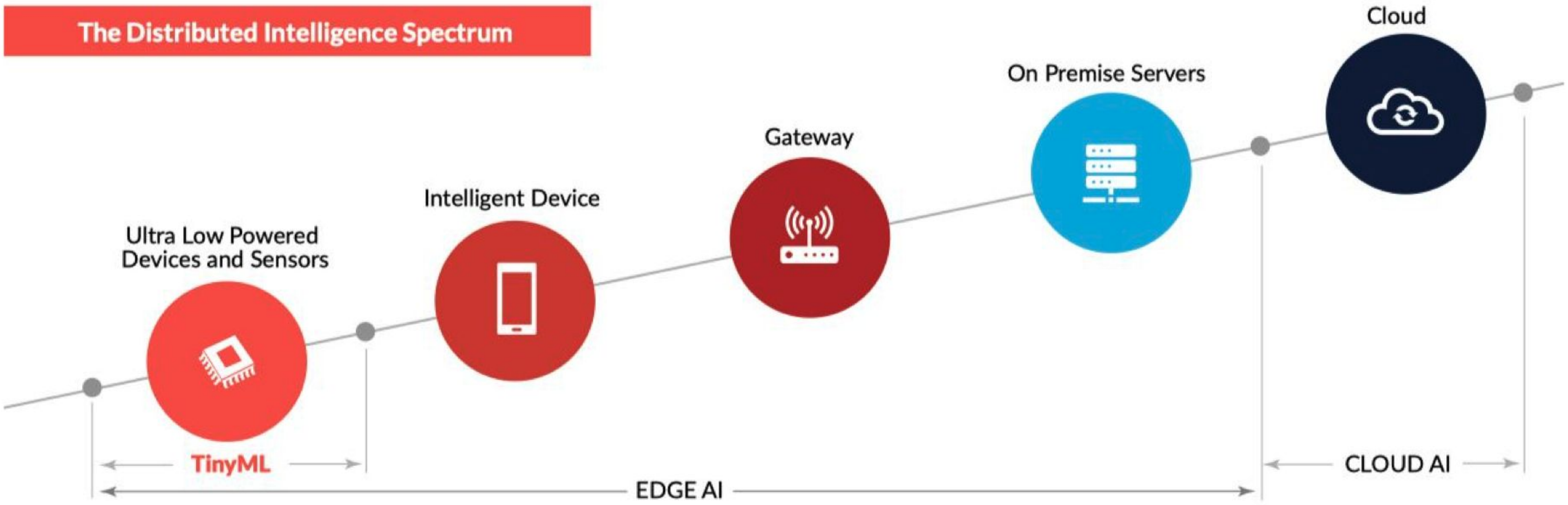
… **Solution ?**

IoT 2.0 * – Edge AI/ML

* Intelligence of Things

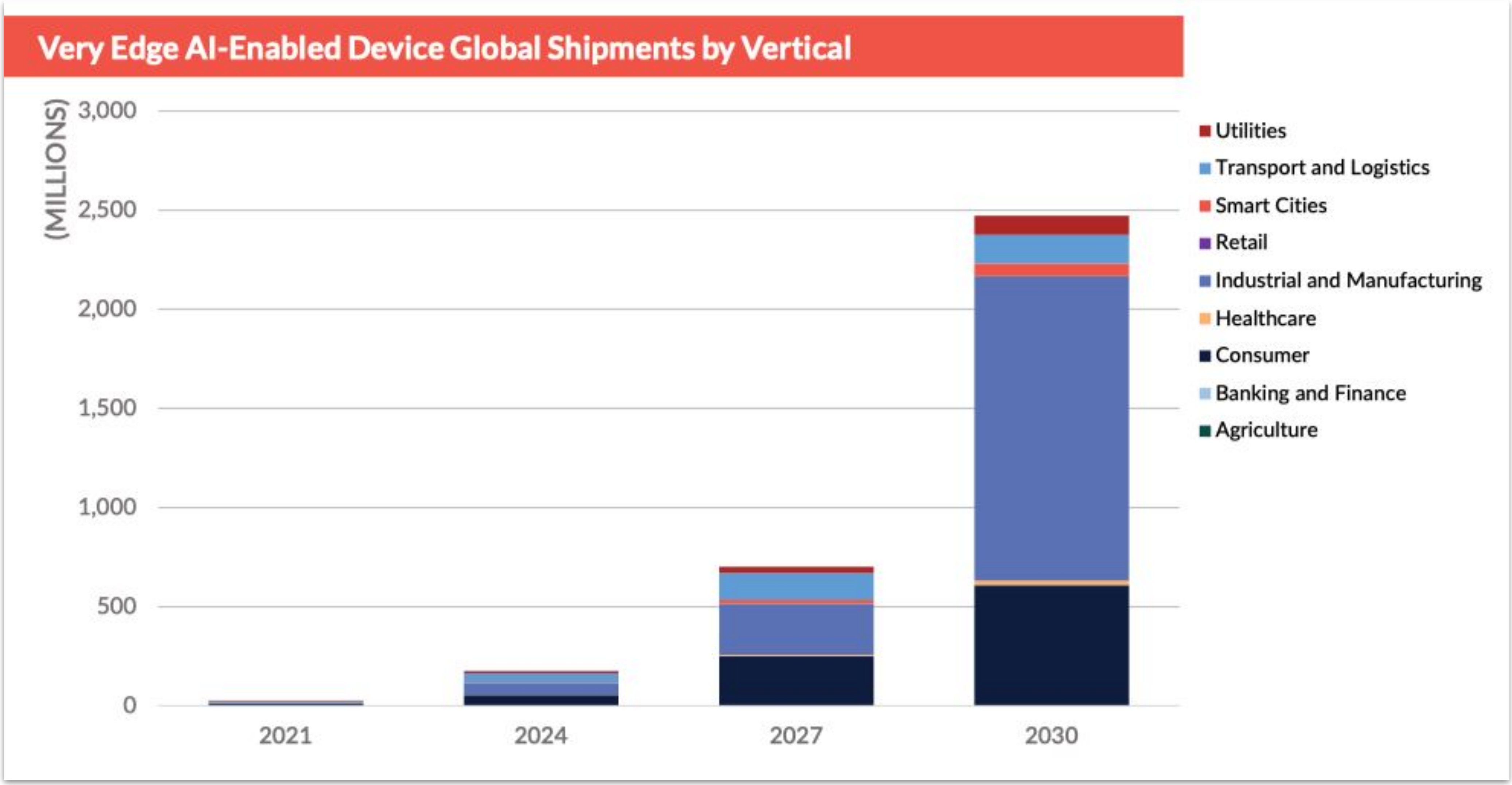


… Solution -> ML goes close to data



Source: ABI Research: TinyML

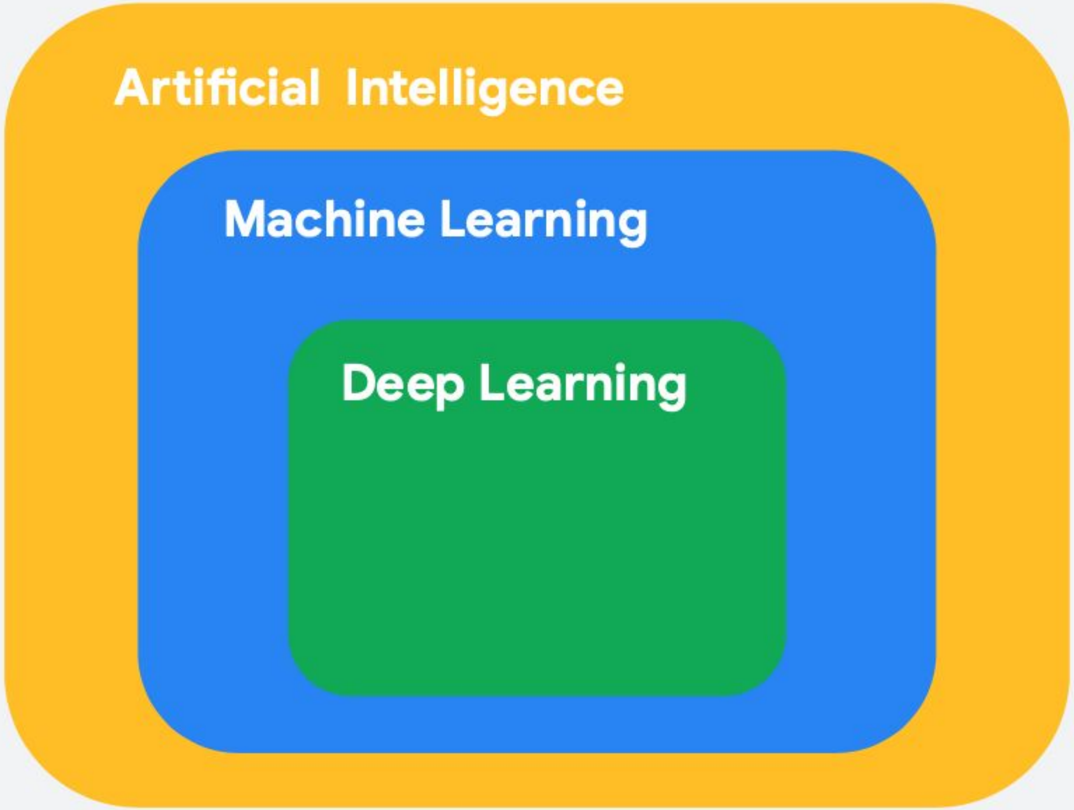
Market Forecast



Source: ABI Research: TinyML

Embedded ML (TinyML)

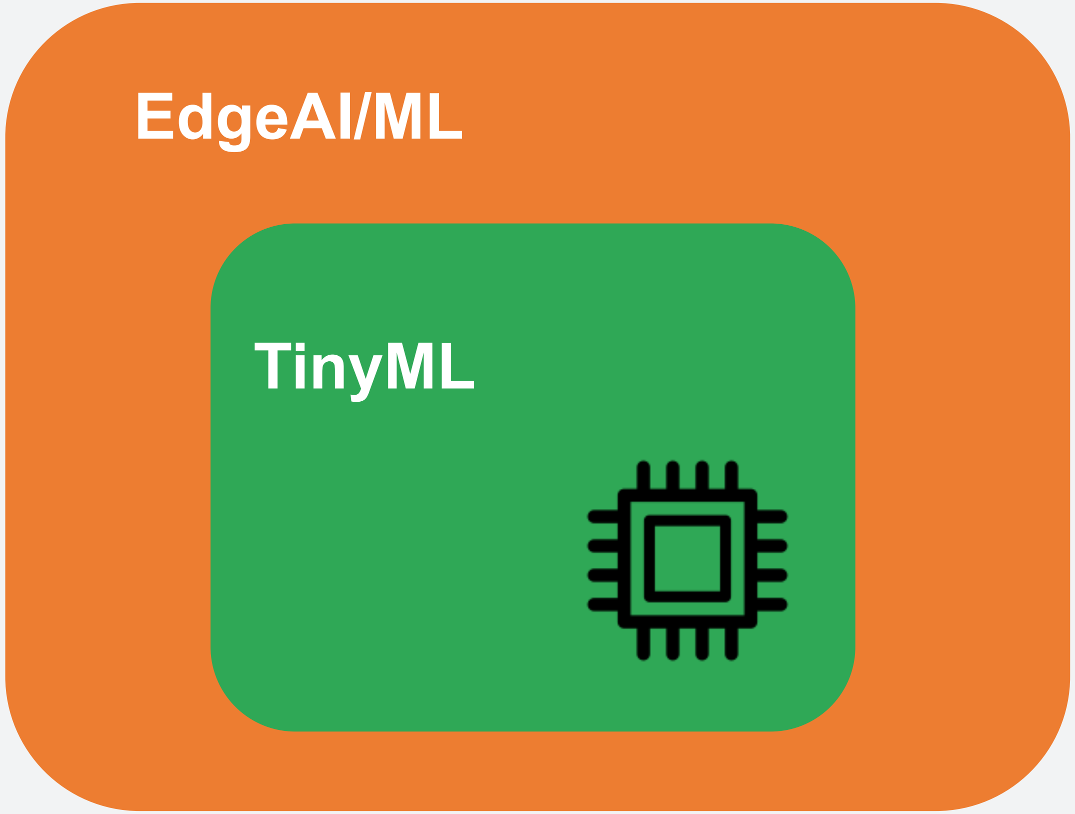
Introduction



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**

19

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

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

20

What is Tiny Machine Learning (TinyML)?

TinyML

Fastest-growing field of ML

Algorithms, hardware, software

On-device sensor analytics

21

What is Tiny Machine Learning (TinyML)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

On-device sensor analytics

Low power consumption

22

What is Tiny Machine Learning (TinyML)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

On-device sensor analytics

Low power consumption

Battery-operated

23

What is Tiny Machine Learning (TinyML)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

On-device sensor analytics

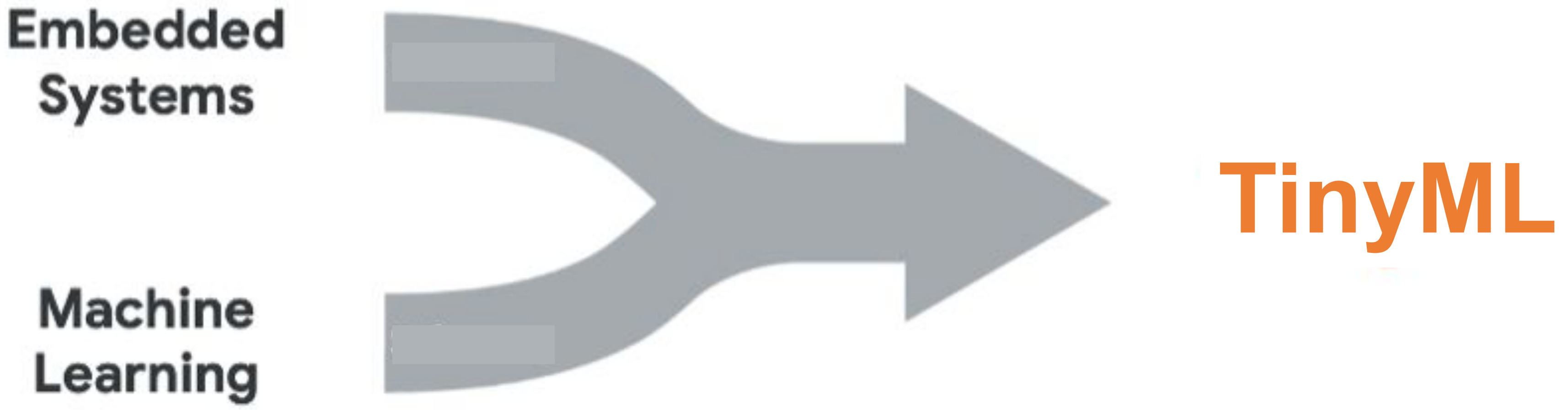
Low power consumption

Battery-operated

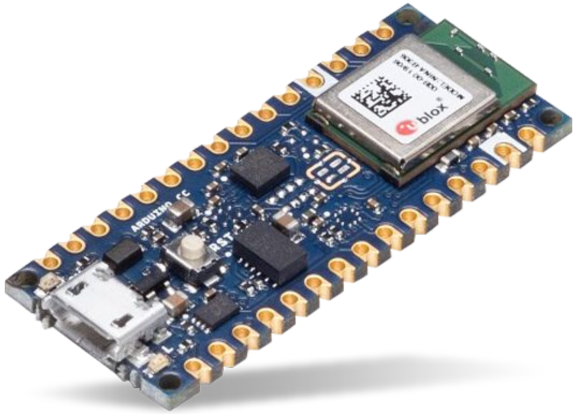
Always-on ML

24

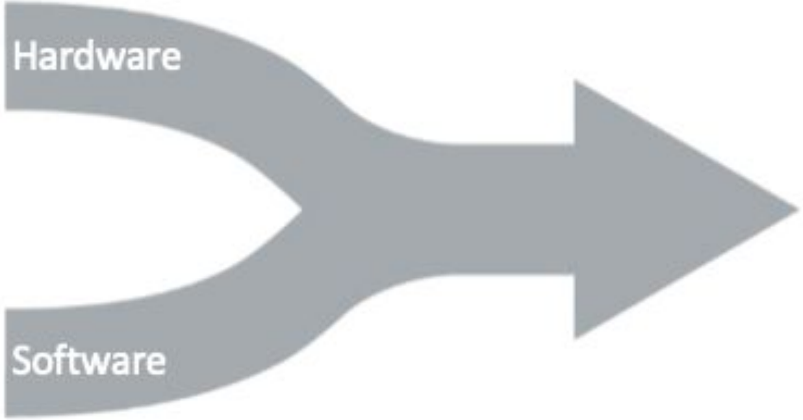
What Makes **TinyML** ?



What Makes TinyML ?

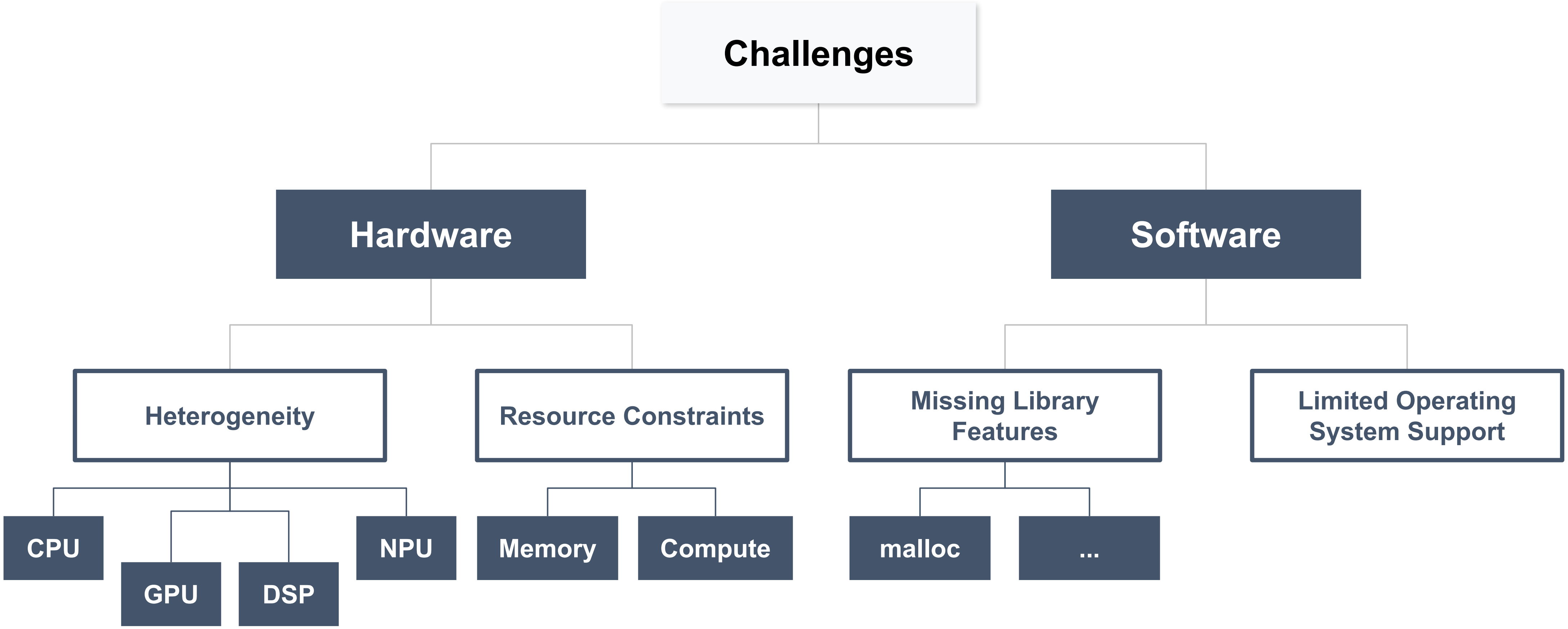


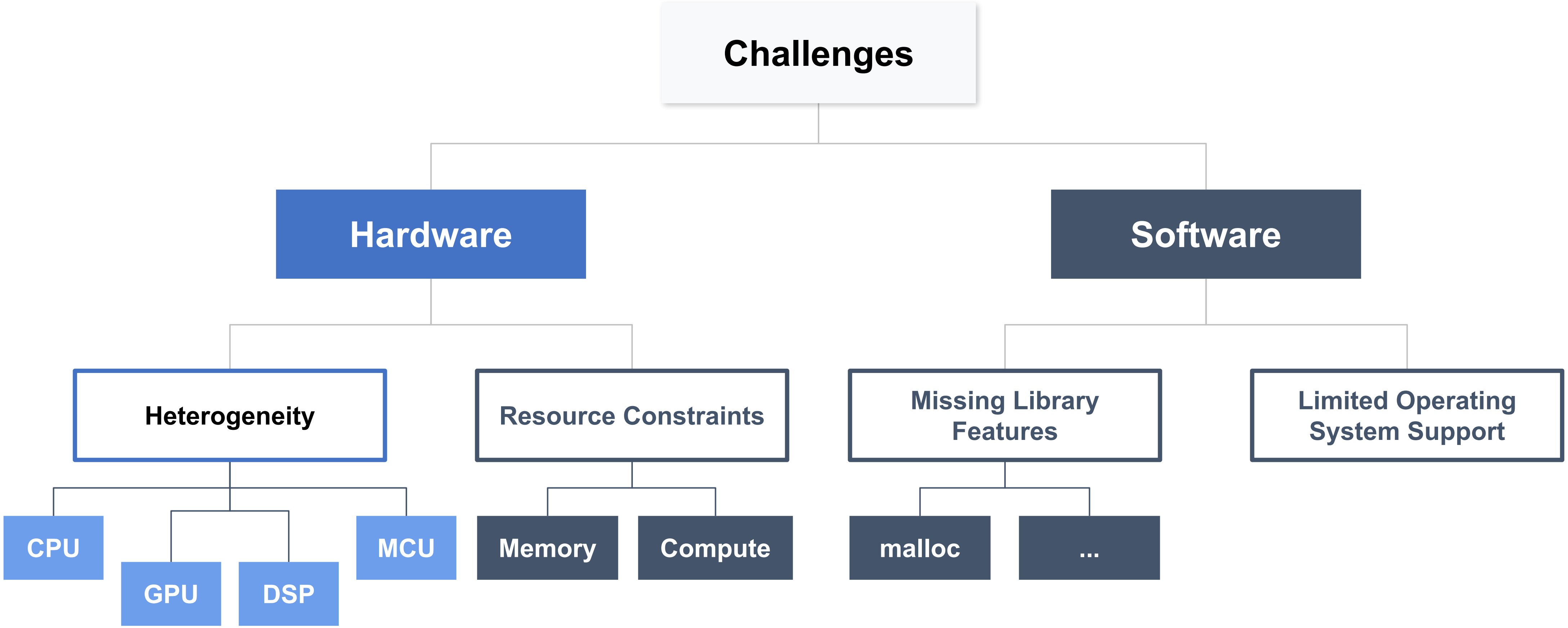
TensorFlow Lite



TinyML

TinyML Challenges


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

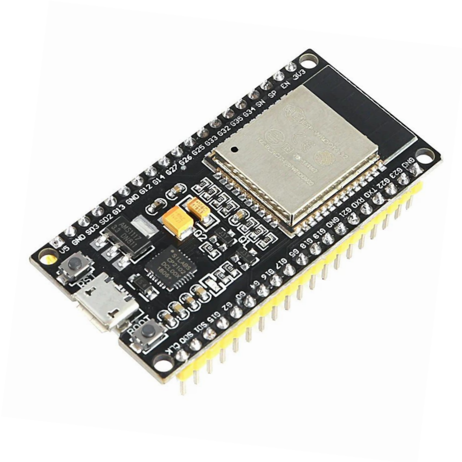

Challenges
Hardware
Software
Heterogeneity
Resource Constraints
Missing Library Features
Limited Operating System Support
CPU
GPU
DSP
MCU
Memory
Compute
malloc
...

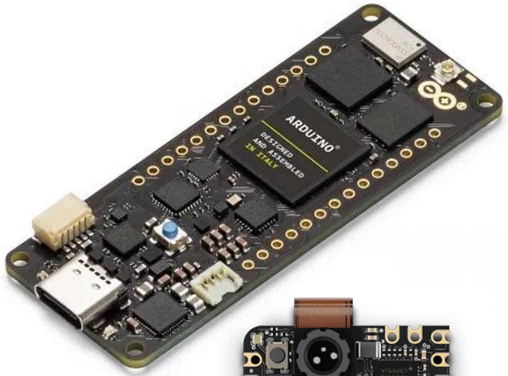
250 Billion

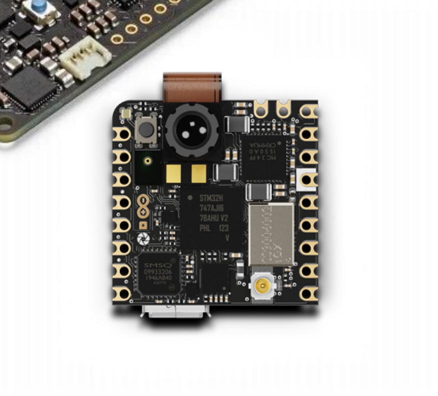
MCUs today

30

Hardware

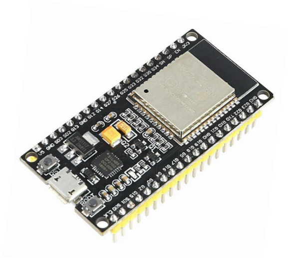


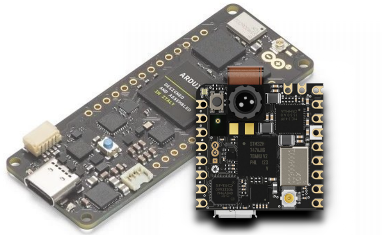




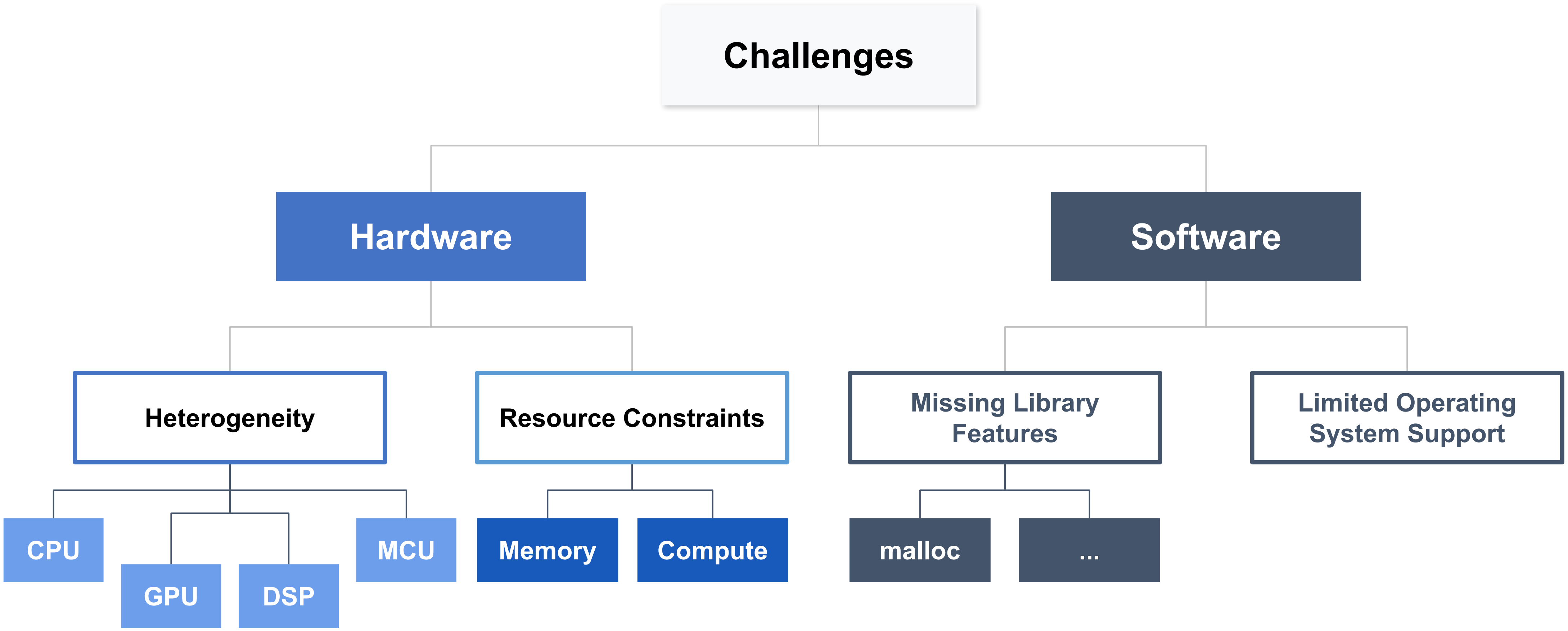
Hardware







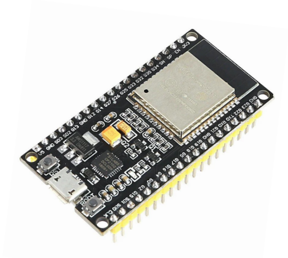
	Raspberry Pico (W)	**Arduino Nano Sense**	**ESP 32**	**Seeed XIAO BLE Sense**	**Arduino Pro**
32Bits CPU	Dual-core Arm Cortex-M0+	Arm Cortex-M4F	Xtensa LX6 Dual Core	Arm Cortex-M4F	Dual Core Arm Cortex M7/M4
CLOCK	133MHz	64MHz	240MHz	64MHz	480/240MHz
RAM	264KB	256KB	520KB	256KB	1MB
ROM	2MB	1MB	2MB	2MB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

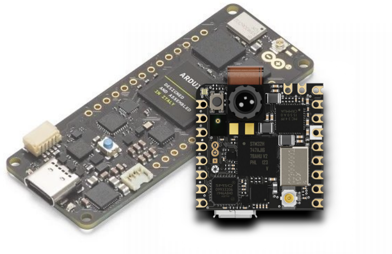

Challenges
Hardware
Software
Heterogeneity
Resource Constraints
Missing Library Features
Limited Operating System Support
CPU
GPU
DSP
MCU
Memory
Compute
malloc
...

33

Hardware

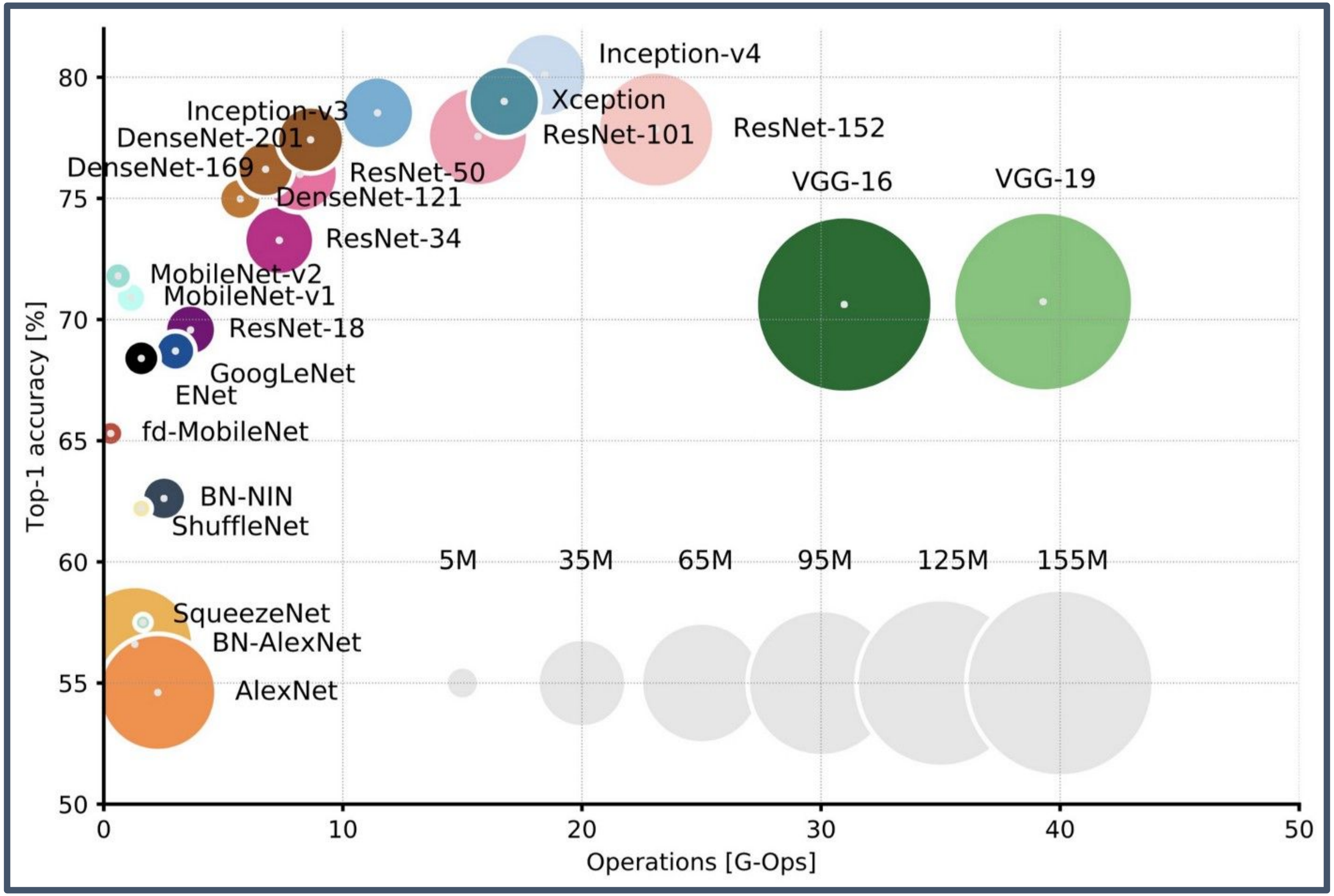






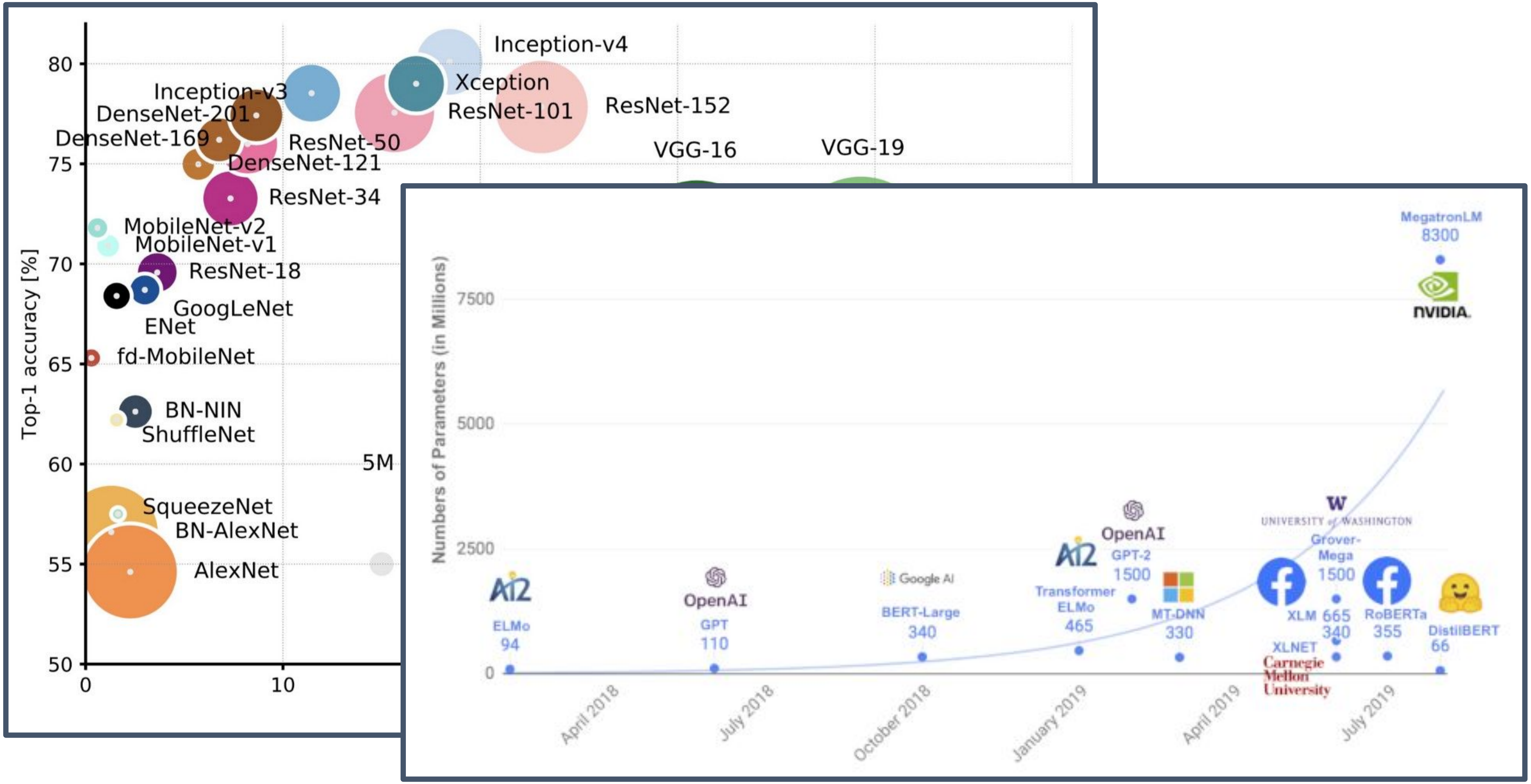
	Raspberry Pico (W)	Arduino Nano Sense	ESP 32	Seeed XIAO BLE Sense	Arduino Pro
32Bits CPU	Dual-core Arm Cortex-M0+	Arm Cortex-M4F	Xtensa LX6 Dual Core	Arm Cortex-M4F	Dual Core Arm Cortex M7/M4
CLOCK	133MHz	64MHz	240MHz	64MHz	480/240MHz
RAM	264KB	256KB	520KB	256KB	1MB
ROM	2MB	1MB	2MB	2MB	2MB
Radio	(Yes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

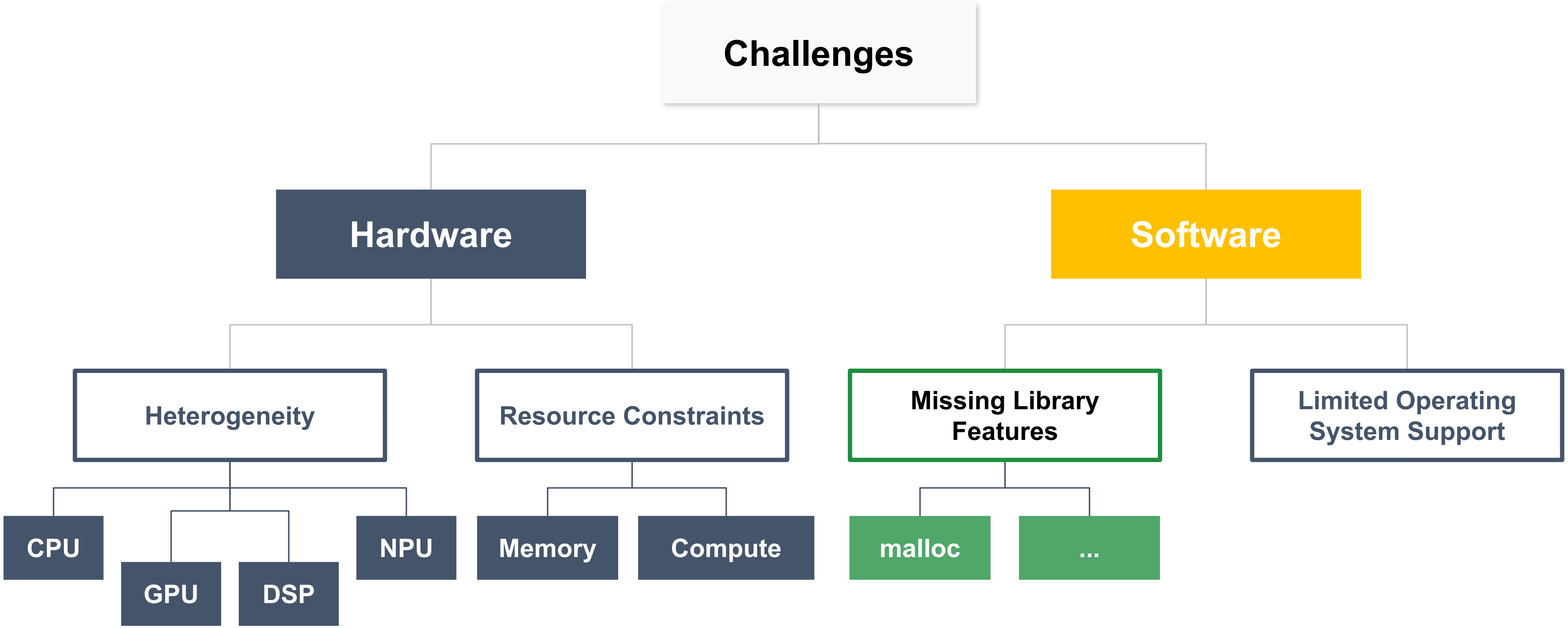
https://arxiv.org/pdf/1910.01108.pdf
https://towardsdatascience.com/neural-network-architectures-156e5bad51ba



35

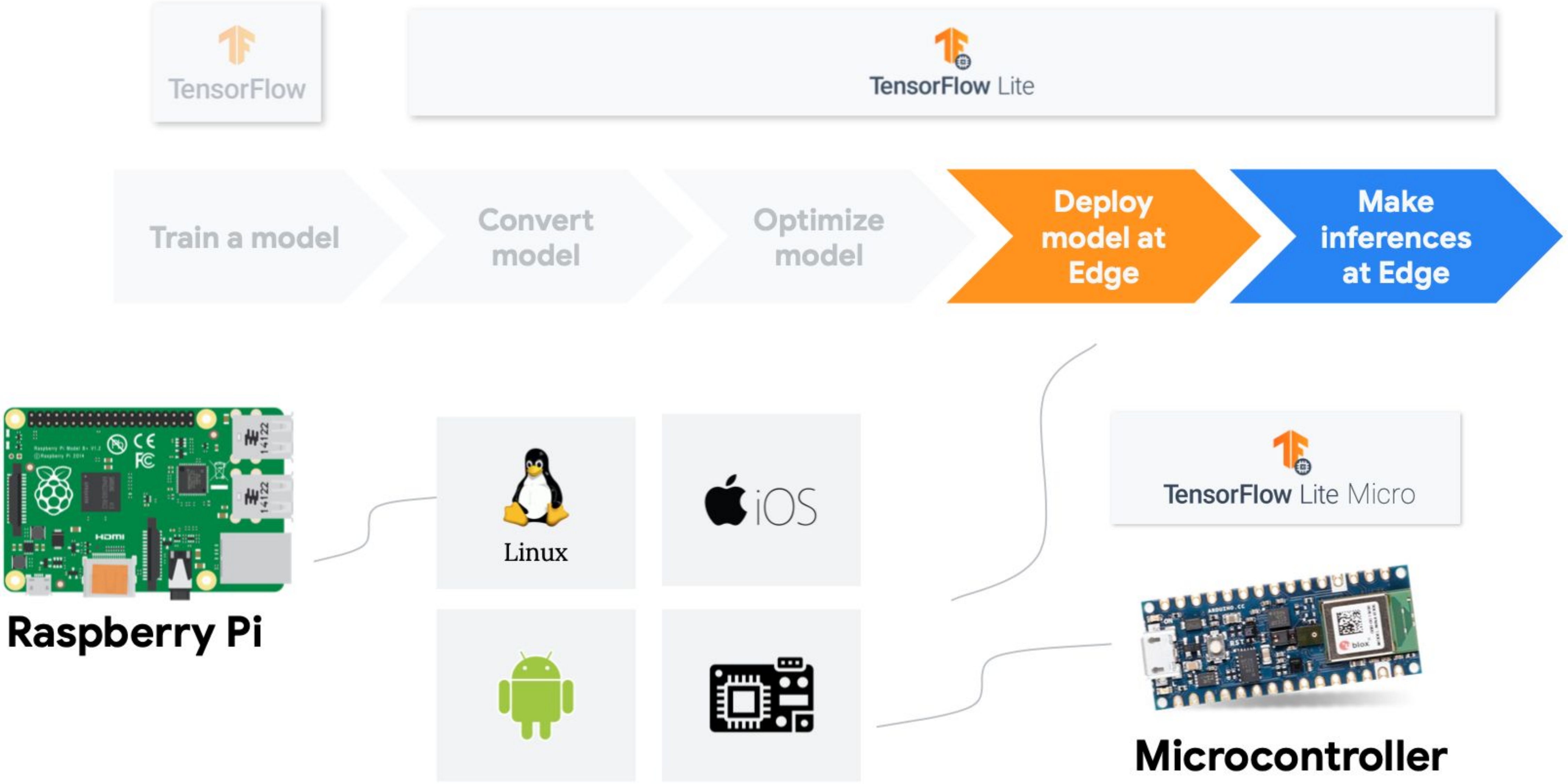
https://arxiv.org/pdf/1910.01108.pdf
https://towardsdatascience.com/neural-network-architectures-156e5bad51ba

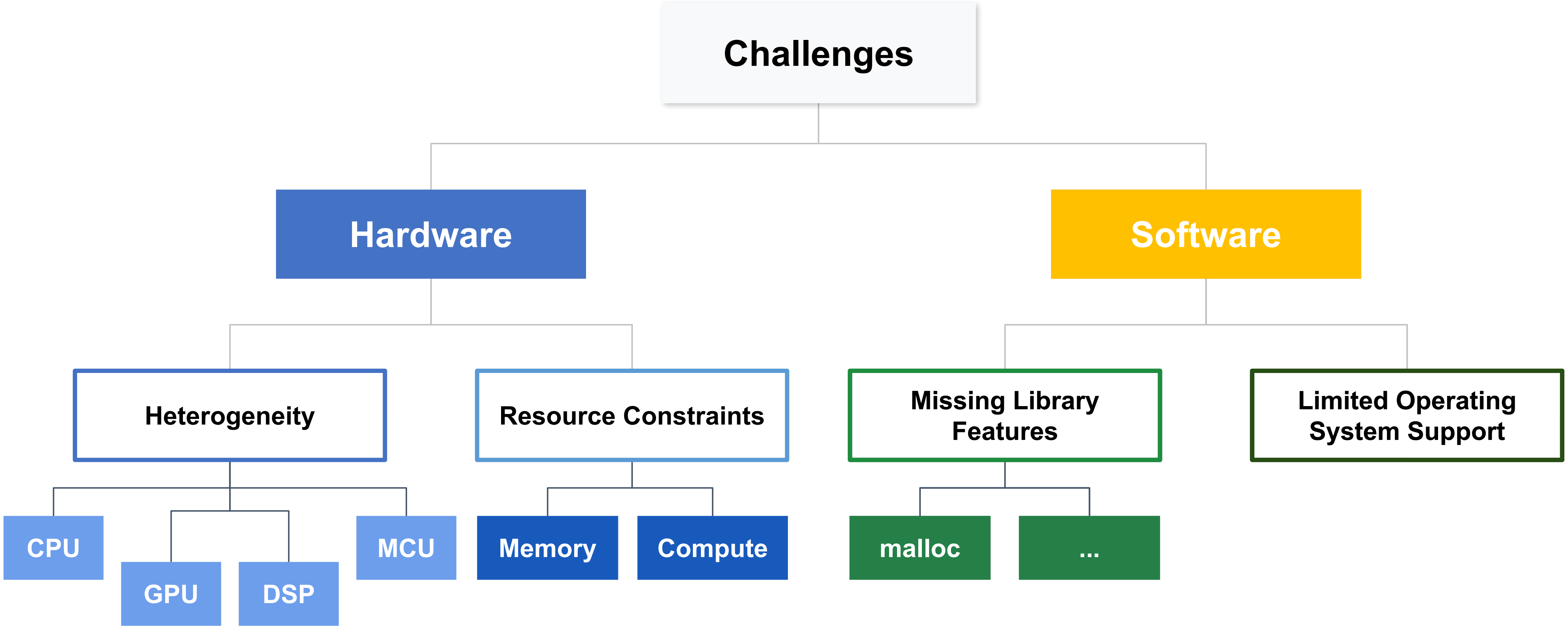



Challenges
Hardware
Software
Heterogeneity
Resource Constraints
Missing Library Features
Limited Operating System Support
CPU
GPU
DSP
NPU
Memory
Compute
malloc
...

37

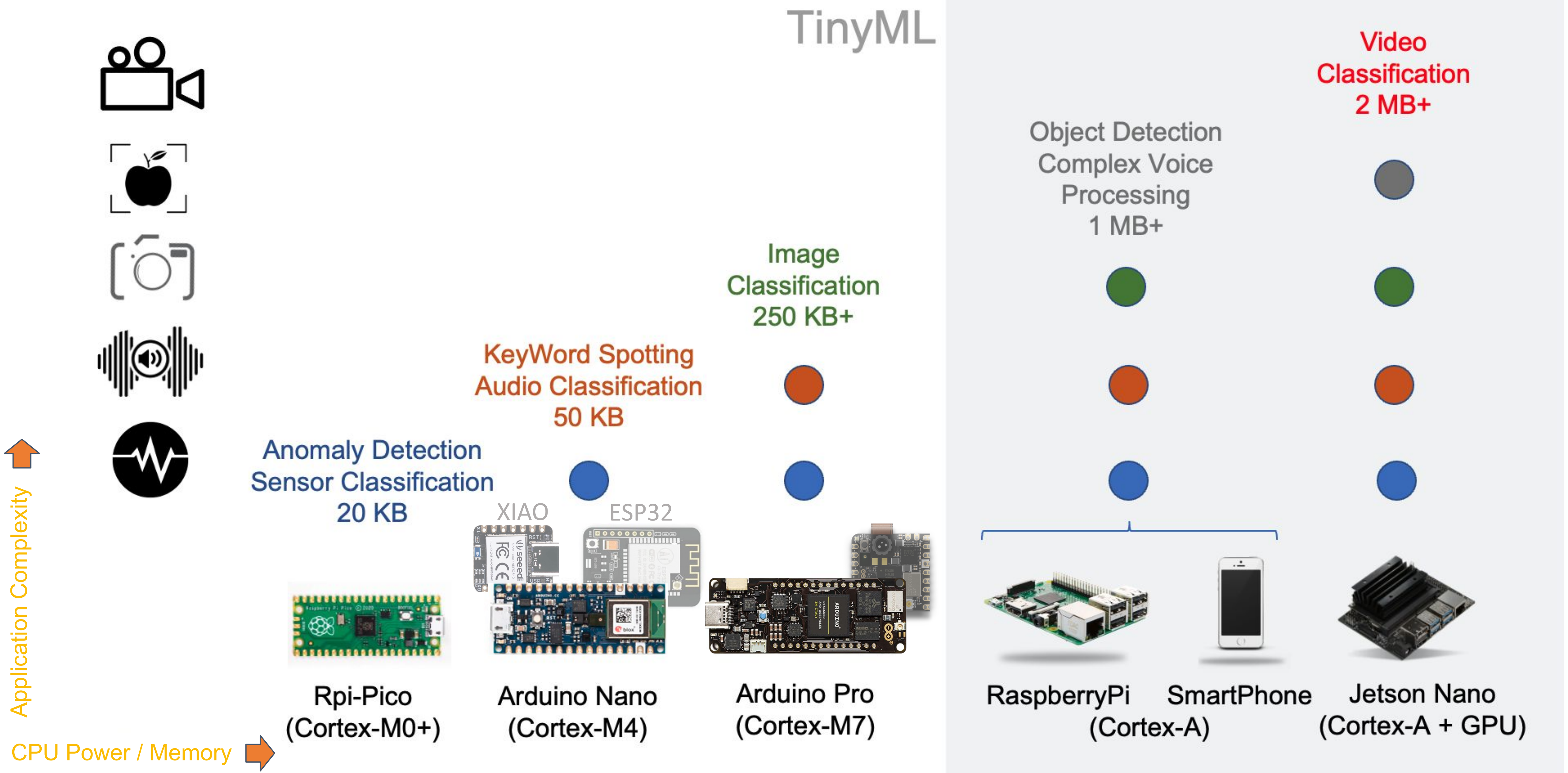
Software


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


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

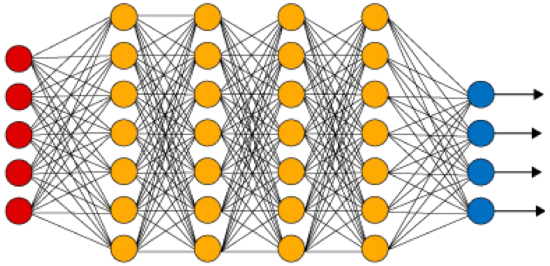
39

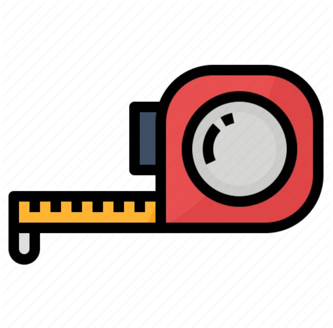
Application Complexity vs. HW



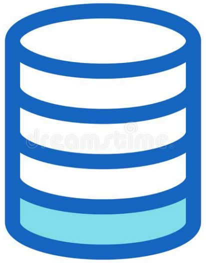
How to Train a ML Model?

Machine Learning Workflow ("What")

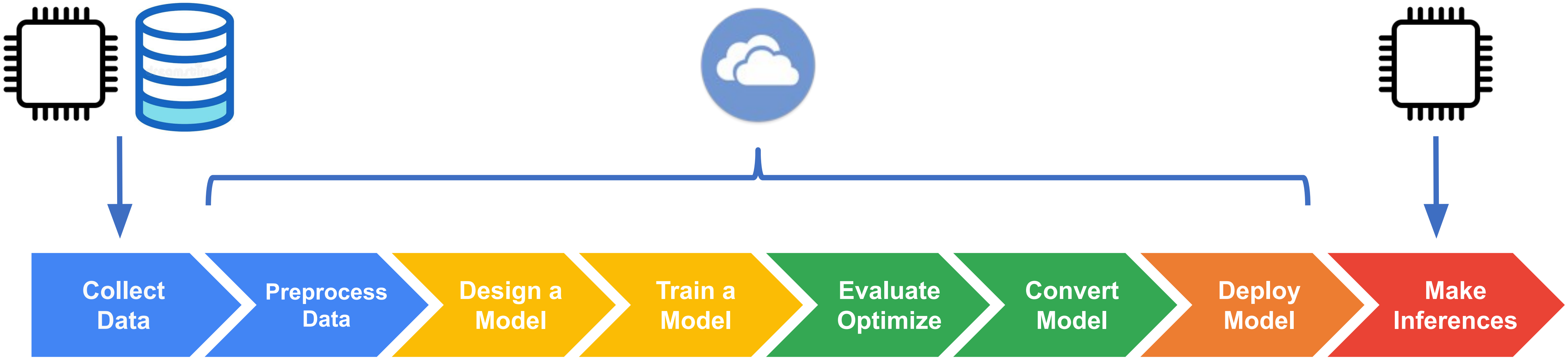




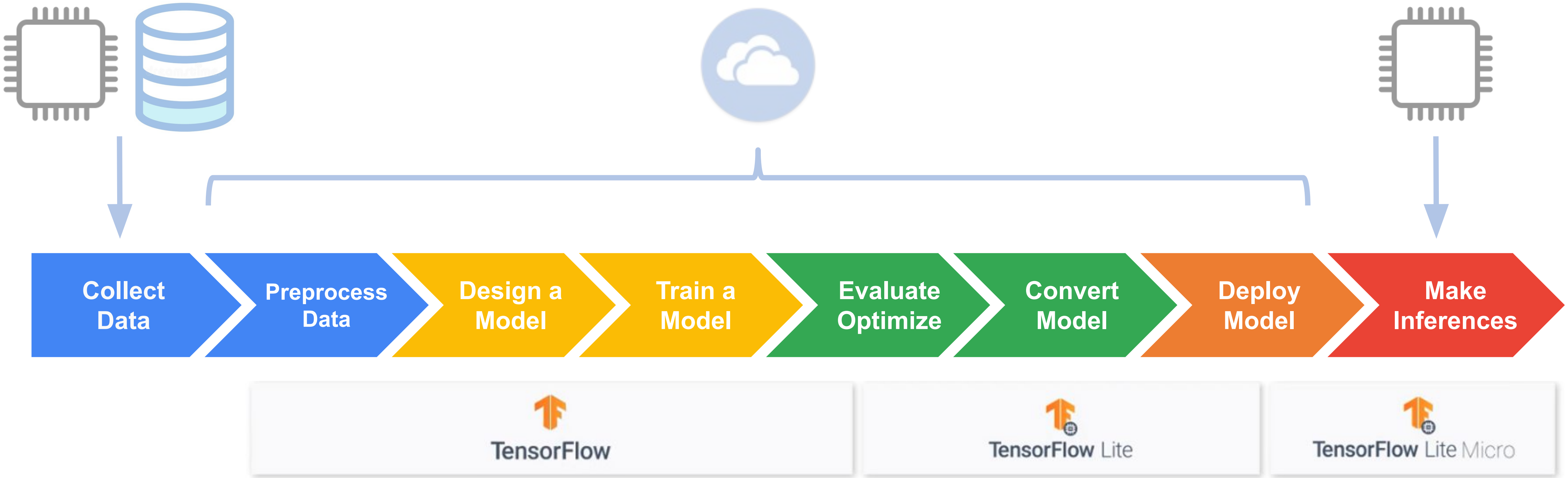
Collect Data → Preprocess Data → Design a Model → Train a Model → Evaluate Optimize → Convert Model → Deploy Model → Make Inferences



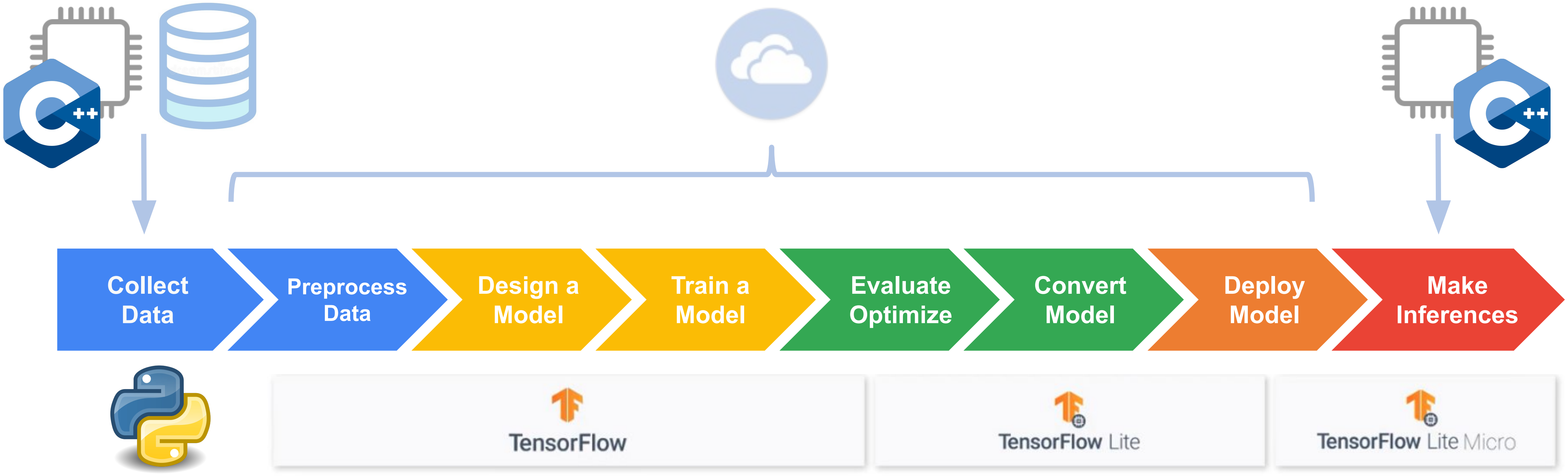
Machine Learning Workflow ("Where")



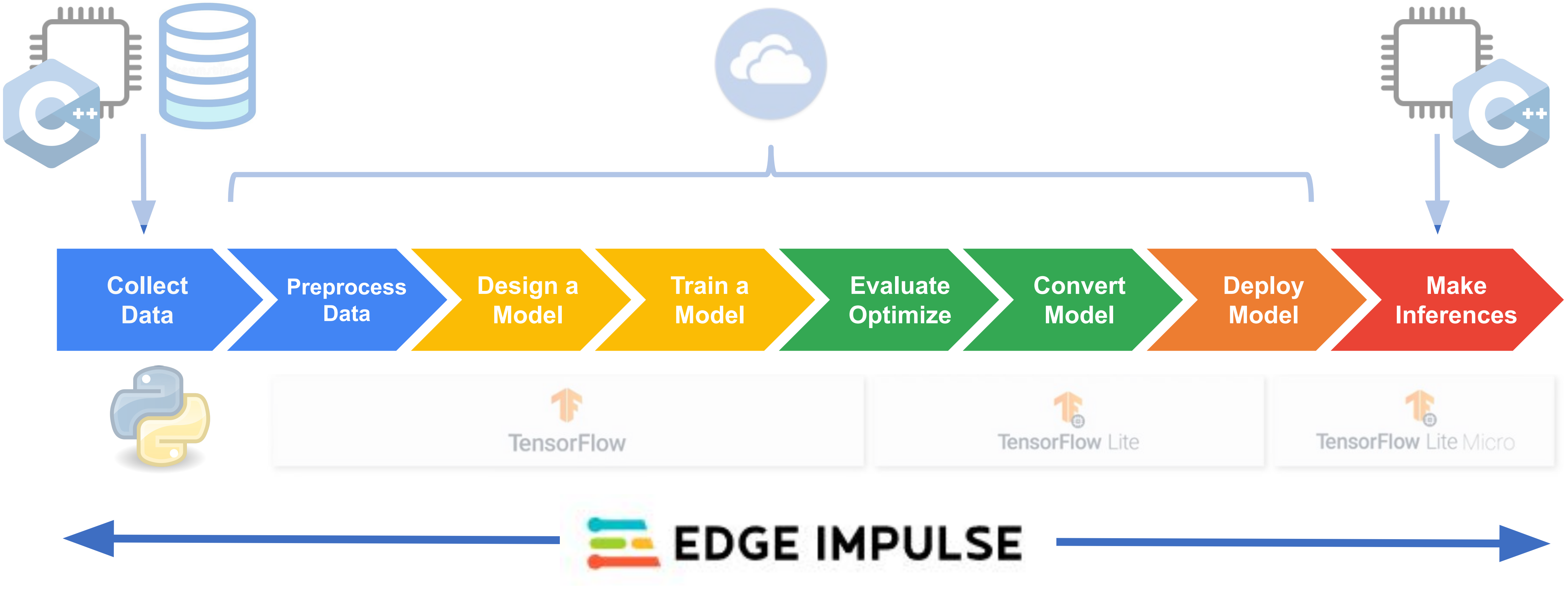
Machine Learning Workflow ("How")



Machine Learning Workflow ("How")

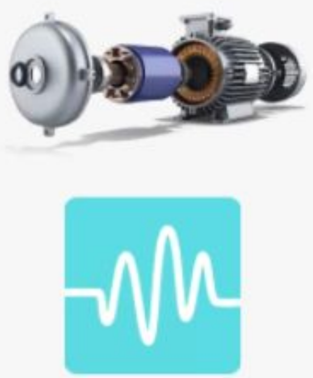


Machine Learning Workflow ("How")



From rule-based engineering to…

Rule-based Code



Sensor

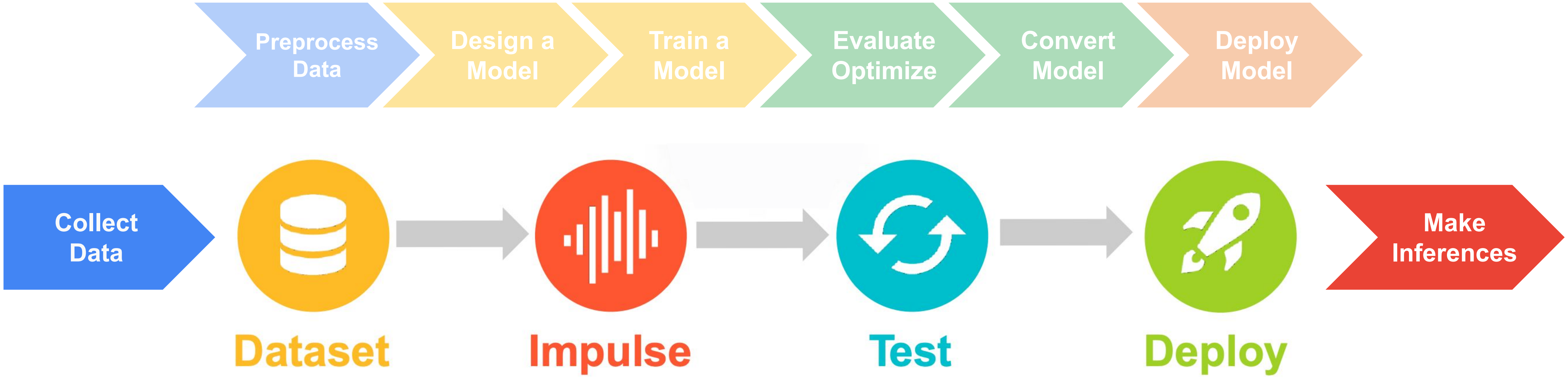
```
int sensorPin = A0;    // select the input pin for the potentiometer
int ledPin = 13;      // select the pin for the LED
int sensorValue = 0;  // variable to store the value coming from the sensor

void setup() {
  // declare the ledPin as an OUTPUT:
  pinMode(ledPin, OUTPUT);
}

void loop() {
  // read the value from the sensor:
  sensorValue = analogRead(sensorPin);
  // turn the ledPin on
  digitalWrite(ledPin, HIGH);
  // stop the program for <sensorValue> milliseconds:
  delay(sensorValue);
  // turn the ledPin off:
  digitalWrite(ledPin, LOW);
  // stop the program for for <sensorValue> milliseconds:
  delay(sensorValue);
}
```

Trial
&
Error

Write More Code

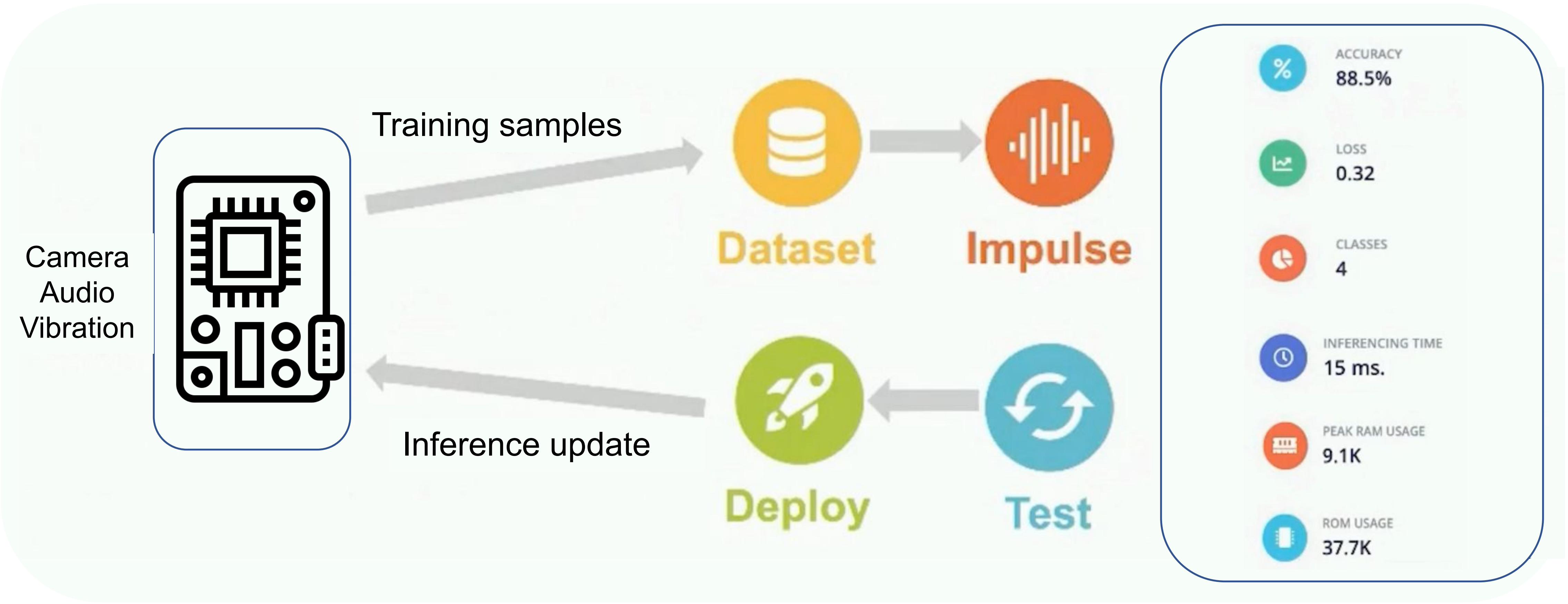
Data-driven engineering

98.7%	0%
0%	100%
0%	0%
0%	0%
0%	0%
0%	0%
0.99	1.00

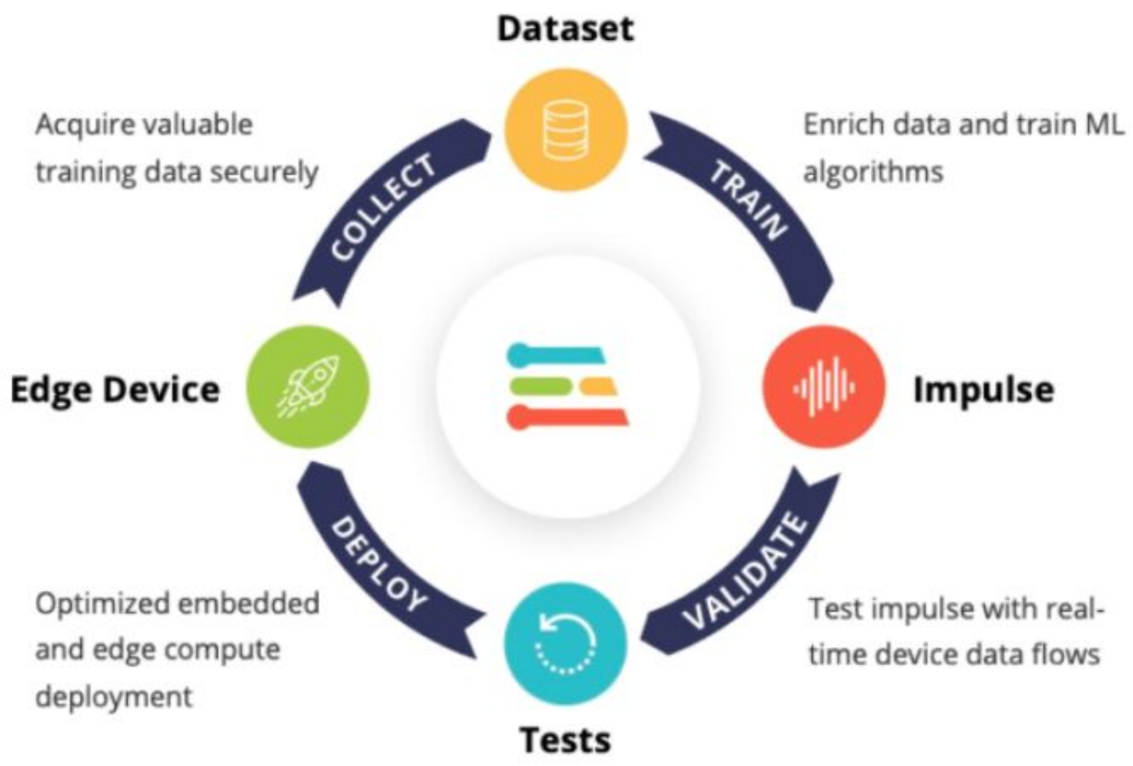


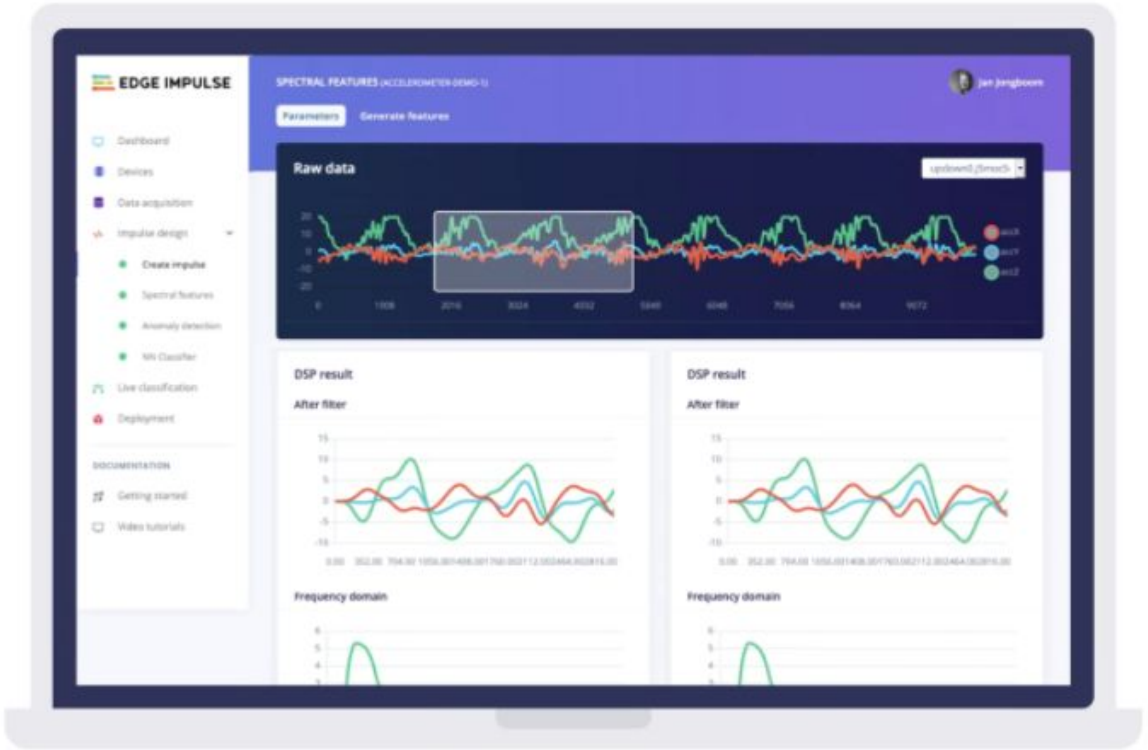
Preprocess Data
Design a Model
Train a Model
Evaluate Optimize
Convert Model
Deploy Model
Collect Data
Dataset
Impulse
Test
Deploy
Make Inferences

Data-driven engineering



EI Studio - Embedded ML platform ("AutoML")

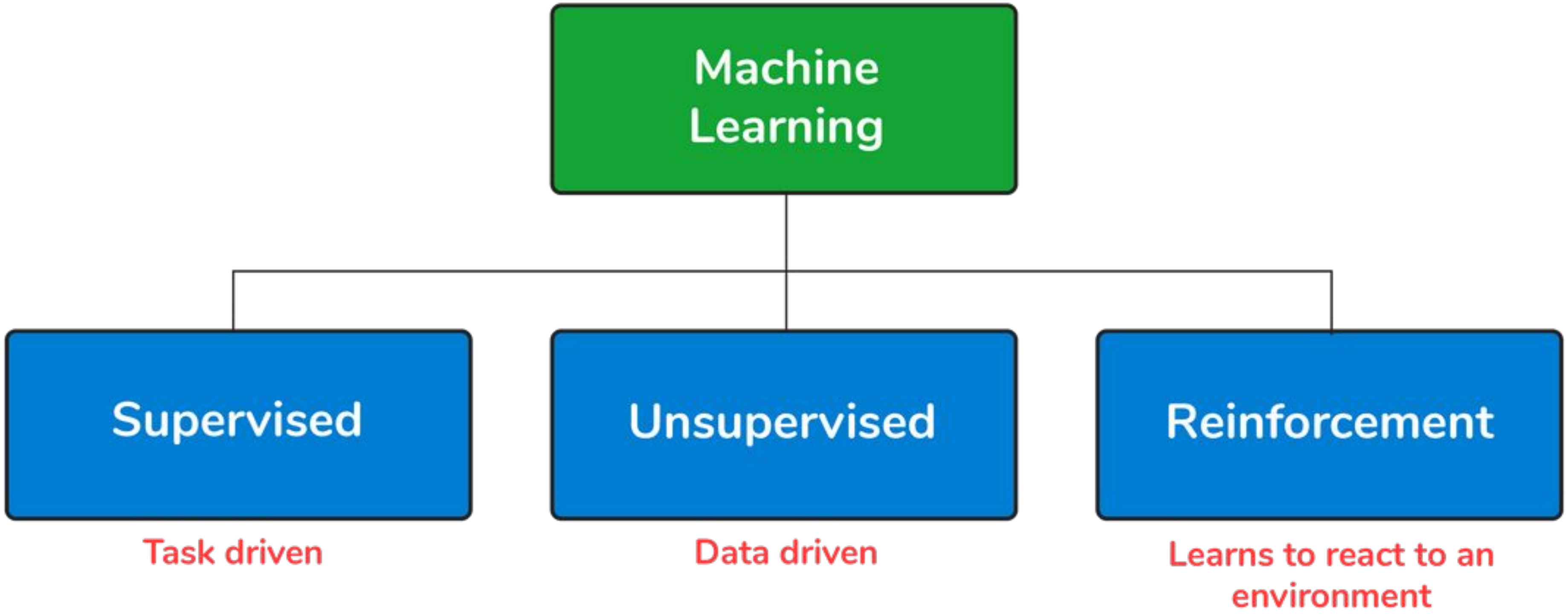




Learn more at http://edgeimpulse.com

TinyML Application

Examples



- Regression
- Classification
- Object Detection

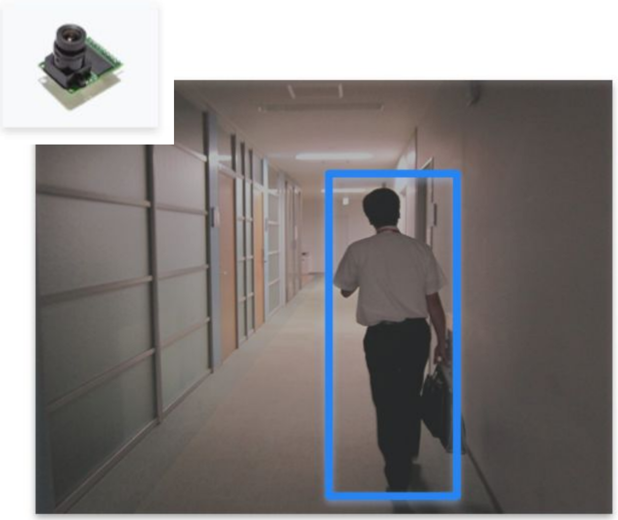
- Anomaly Detection

- Autonomous Navigation

Sound

Vibration

Vision

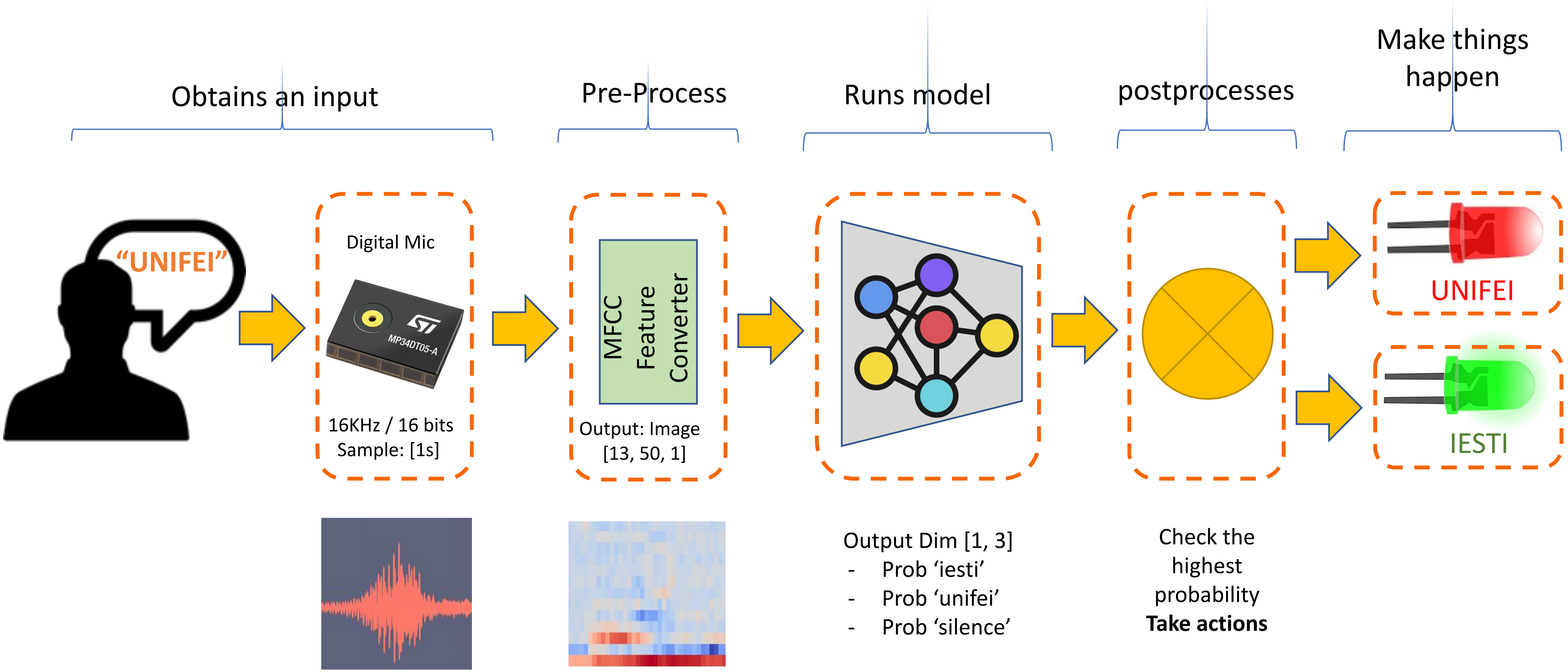


Sound

Vibration

Vision

KeyWord Spotting (KWS) - **Inference**



https://youtu.be/XnFYz-RSNe8

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

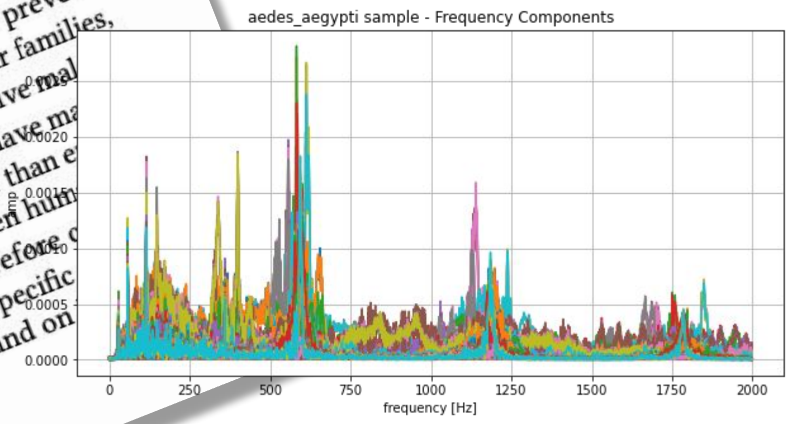
ABSTRACT

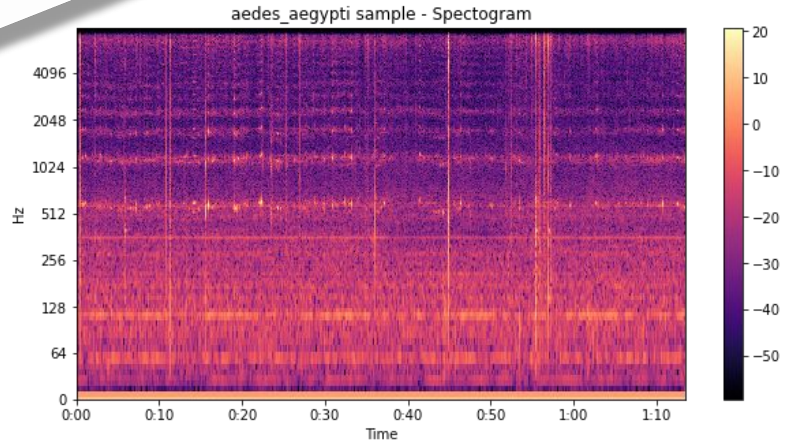
Every year more than one billion people are infected and more than one million people die from vector-borne diseases including malaria, dengue, zika and chikungunya. Mosquitoes are the best known disease vector and are geographically spread worldwide. It is important to raise awareness of mosquito proliferation 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.





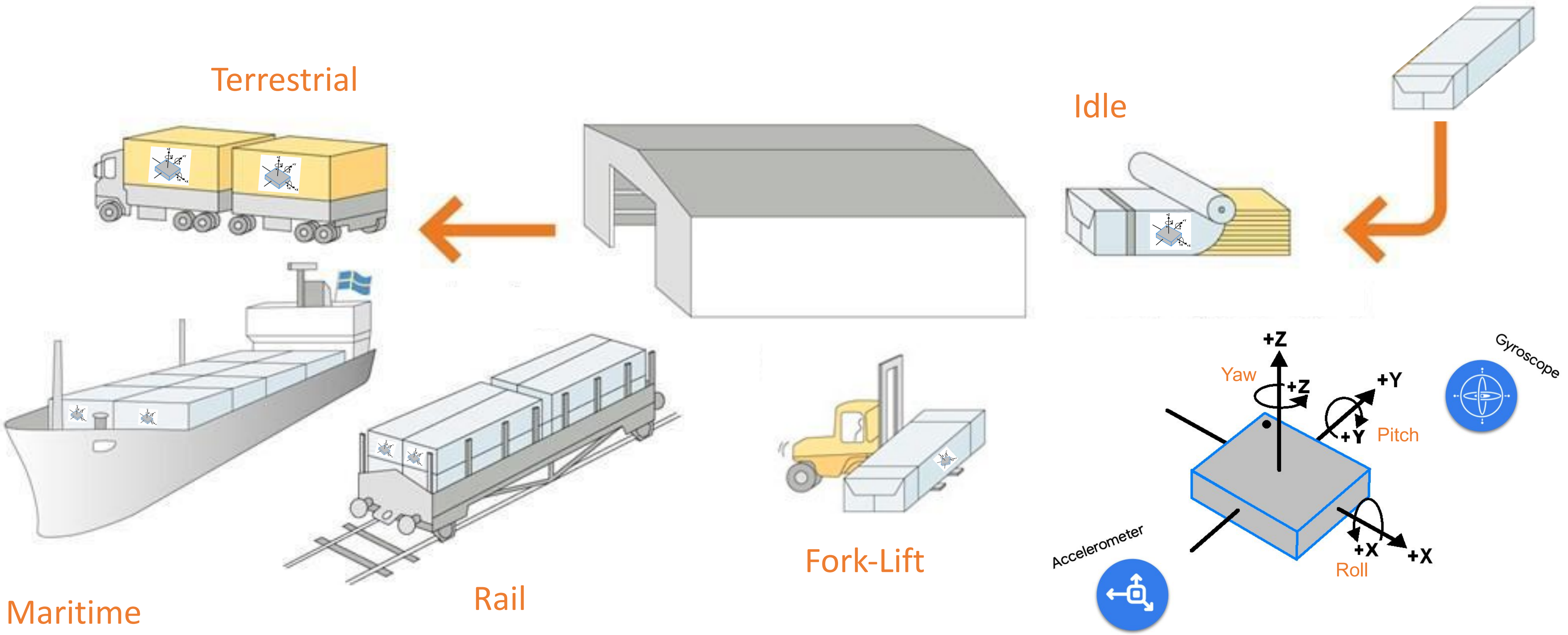
https://github.com/Mjrovai/wingbeat-mosquito-tinyml

Sound

Vibration

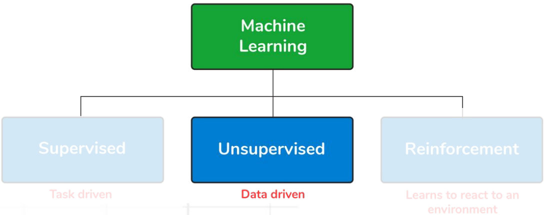
Vision

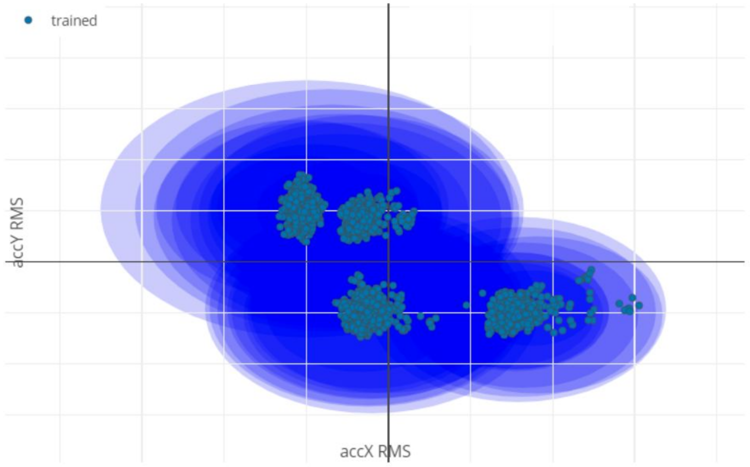
Mechanical Stresses in Transport



[ICTP SciTyniML 21 - Hands on Embedded ML - Motion/Anomaly Detection and Scientific Applications]

Industrial – Anomaly Detection





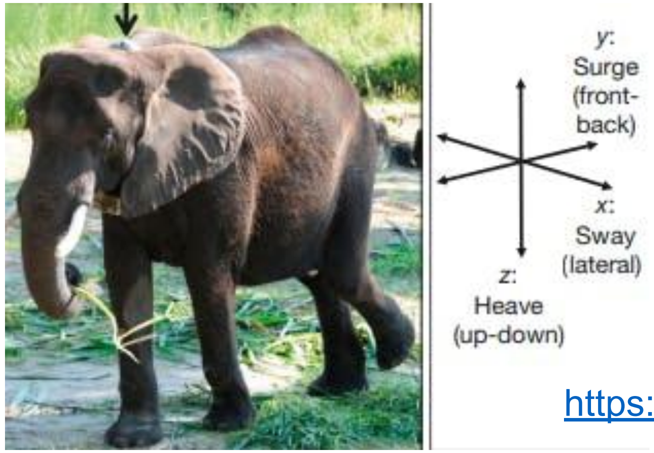
[IESTI01 2021.2 - Final Group Project: Bearing Failure Detection]

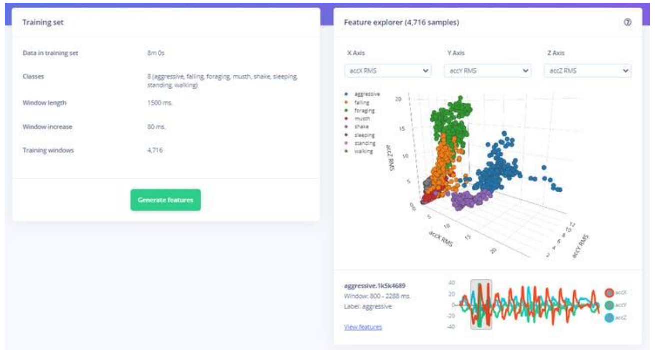
Predict and classify common Elephant behavior

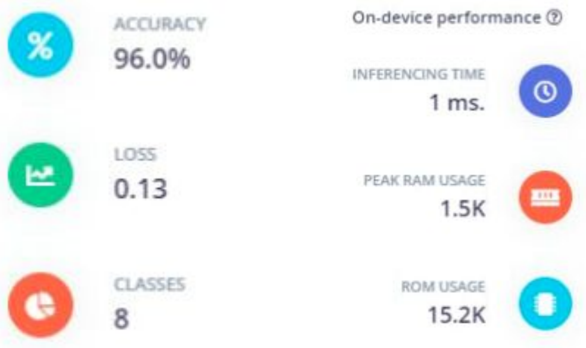
Aggressive

Standing

Sleeping





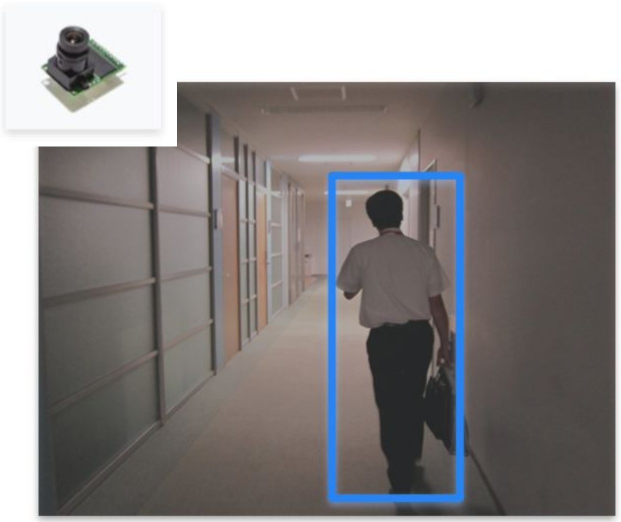


https://www.hackster.io/dhruvsheth_/eletect-tinyml-and-iot-based-smart-wildlife-tracker-c03e5a

Sound

Vibration

Vision



Computer Vision Main Types

Image Classification

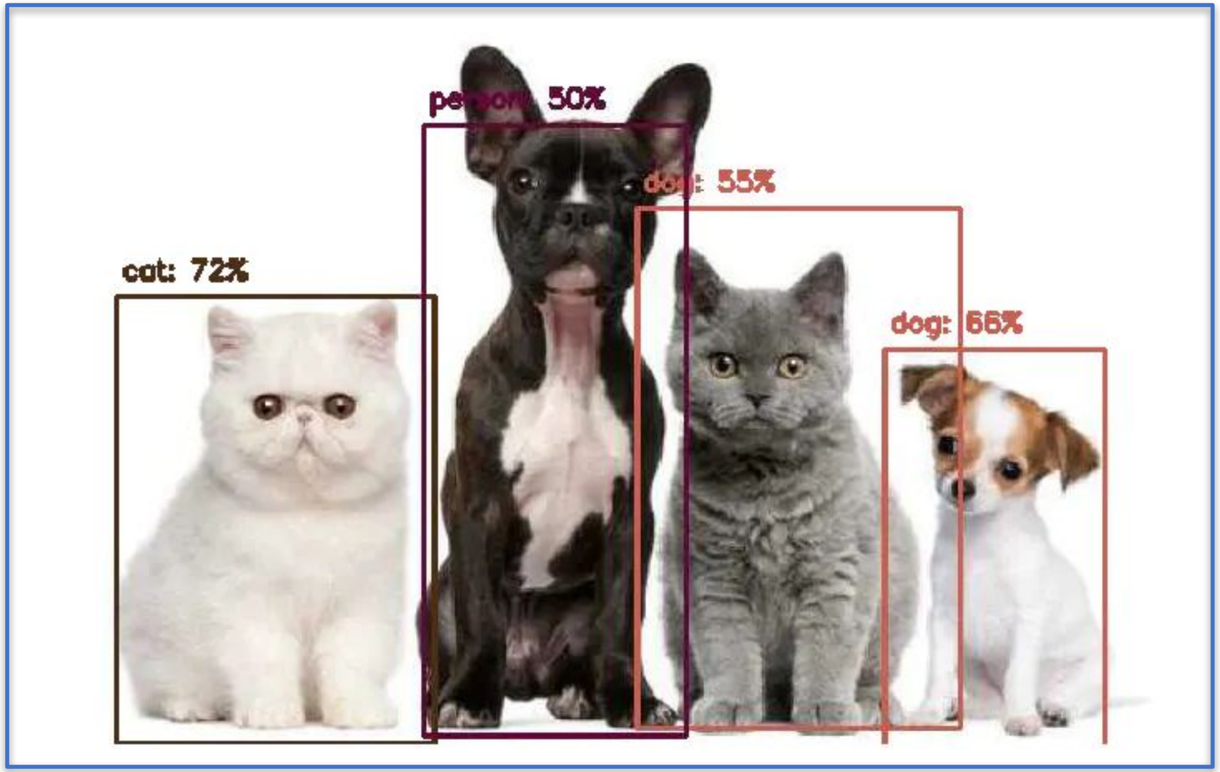
(**Multi-Class** Classification)

Cat: 70%

Dog: 80%

Object Detection

Multi-Label Classification + Object **Localization**



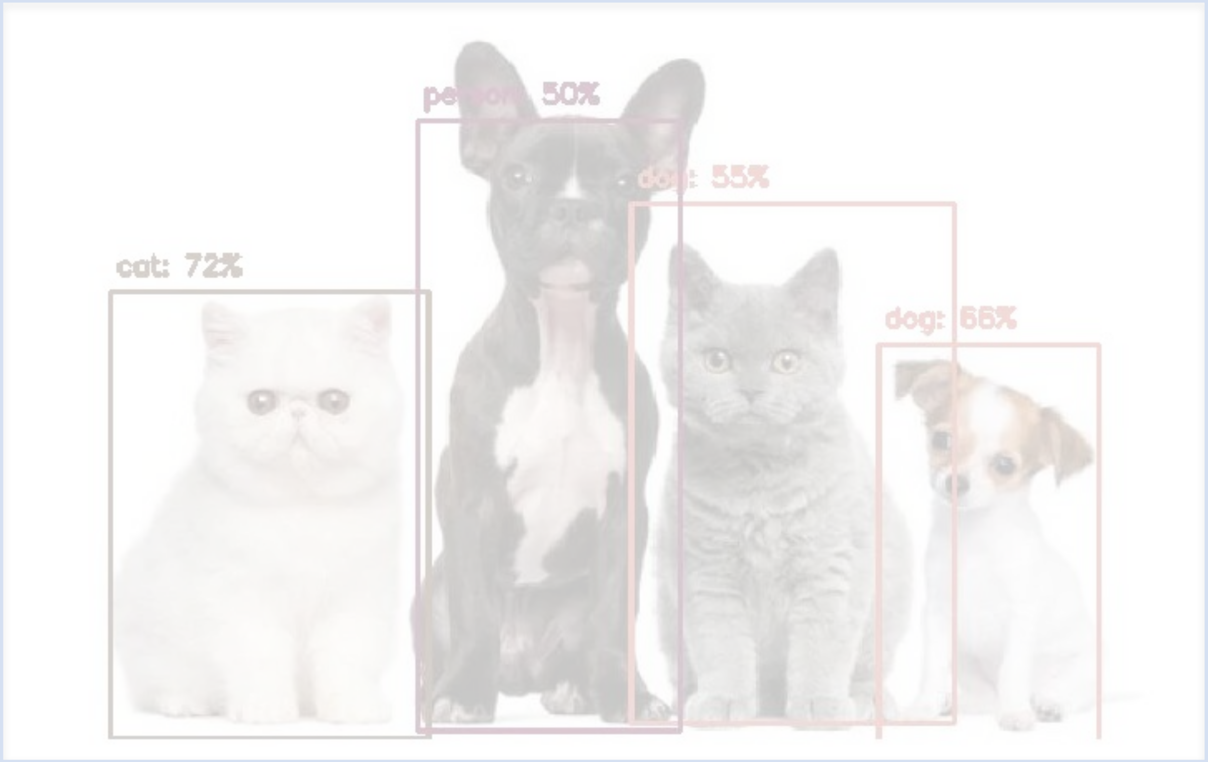
Computer Vision Main Types

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

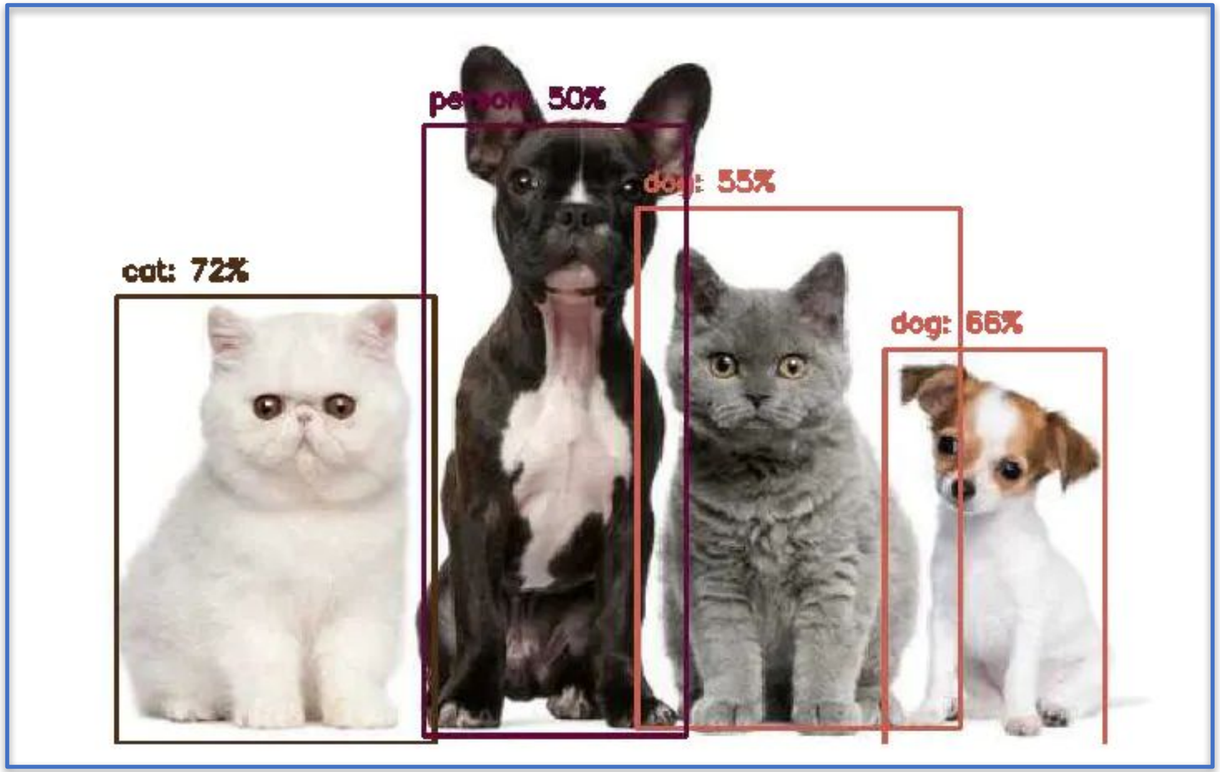
(**Multi-Class** Classification)

Cat: 70%

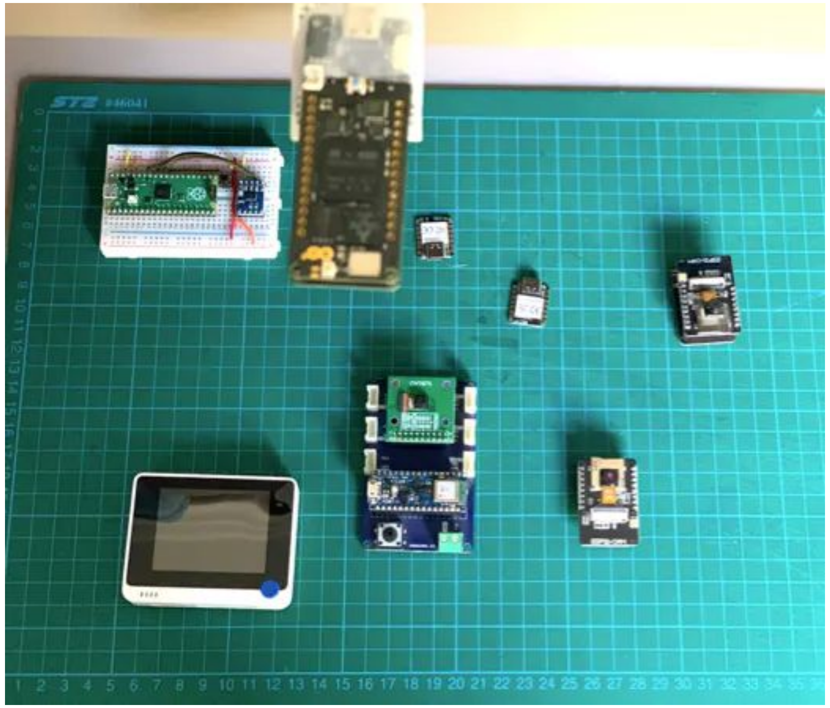
Dog: 80%

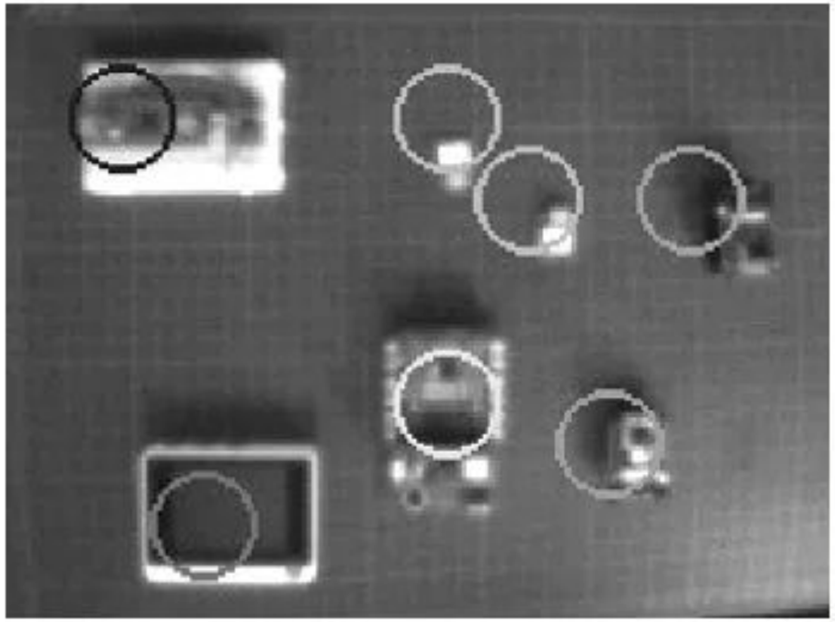
Object Detection

Multi-Label Classification + Object **Localization**



Detecting Objects using TinyML (FOMO)

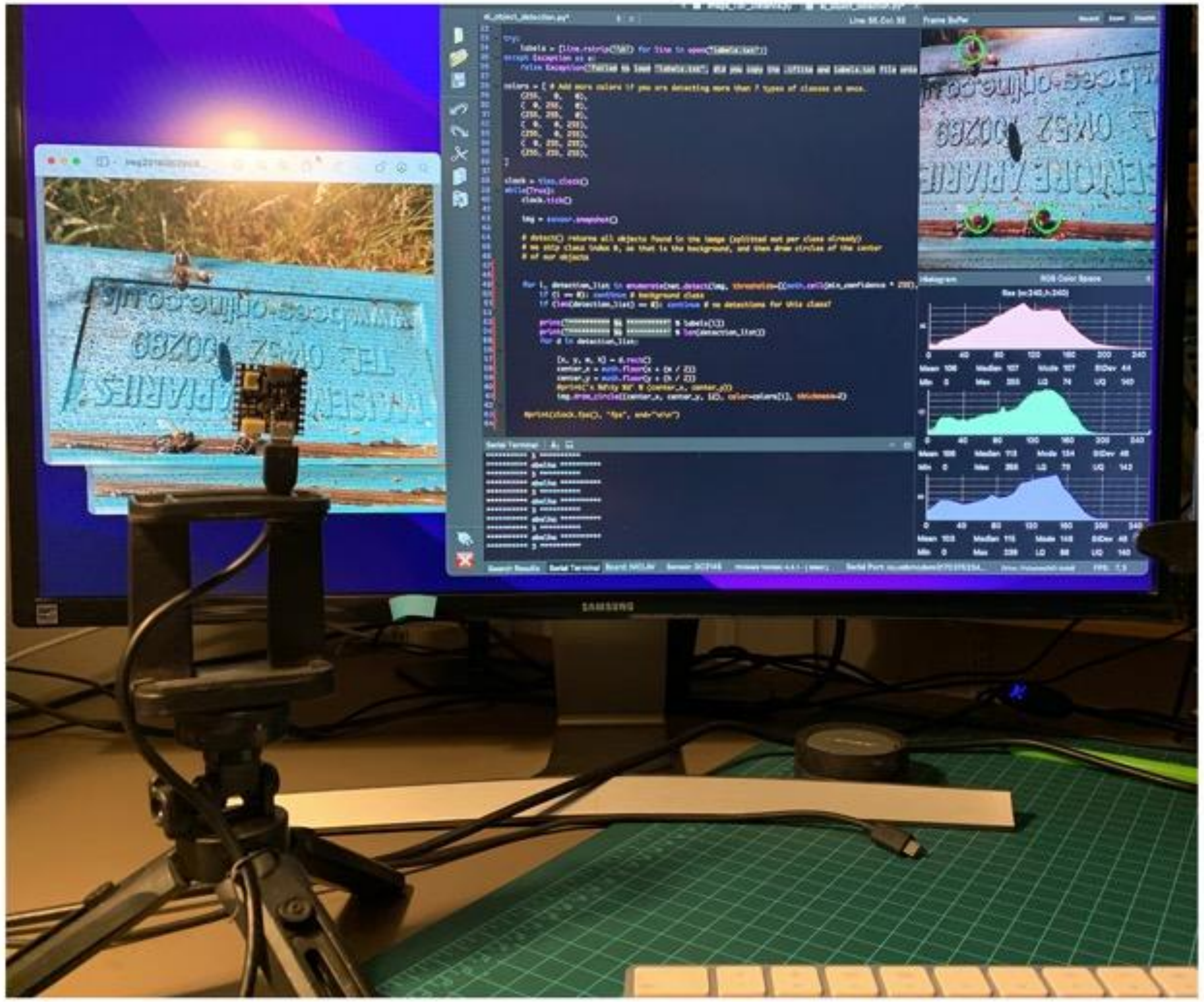


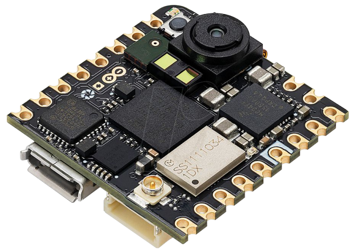


```
********** espcam **********
x 70    y 150
x 130   y 170
********** nano **********
x 70    y 110
********** pico **********
x 150   y 30
********** wio **********
x 50    y 50
********** xiao **********
x 150   y 110
x 130   y 130
6.97512 fps
```

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

Detecting Objects using TinyML (FOMO)





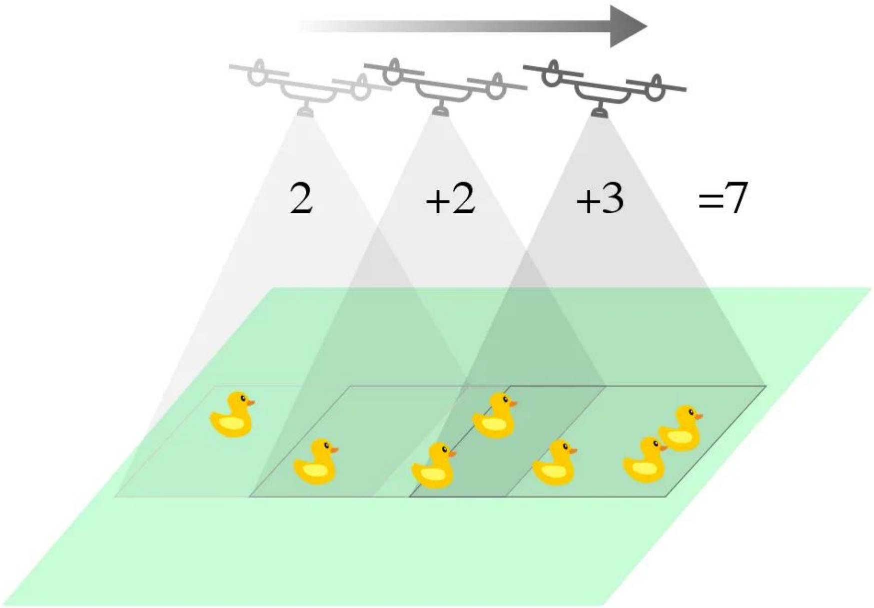
OpenMV

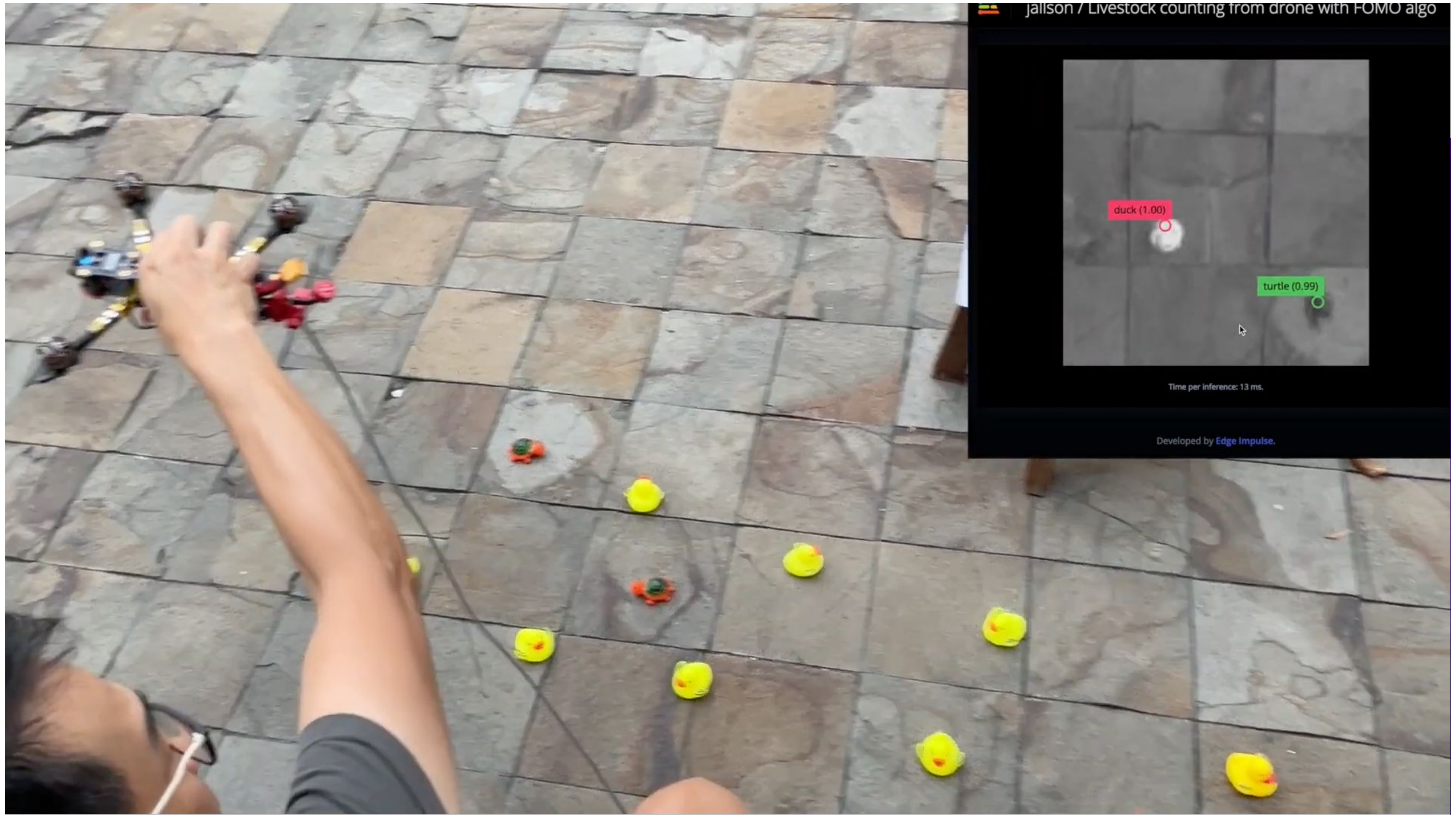
MicroPython

EDGE IMPULSE

https://youtu.be/MYuc3QISquw

Livestock / Wildlife Counting from Drone with FOMO



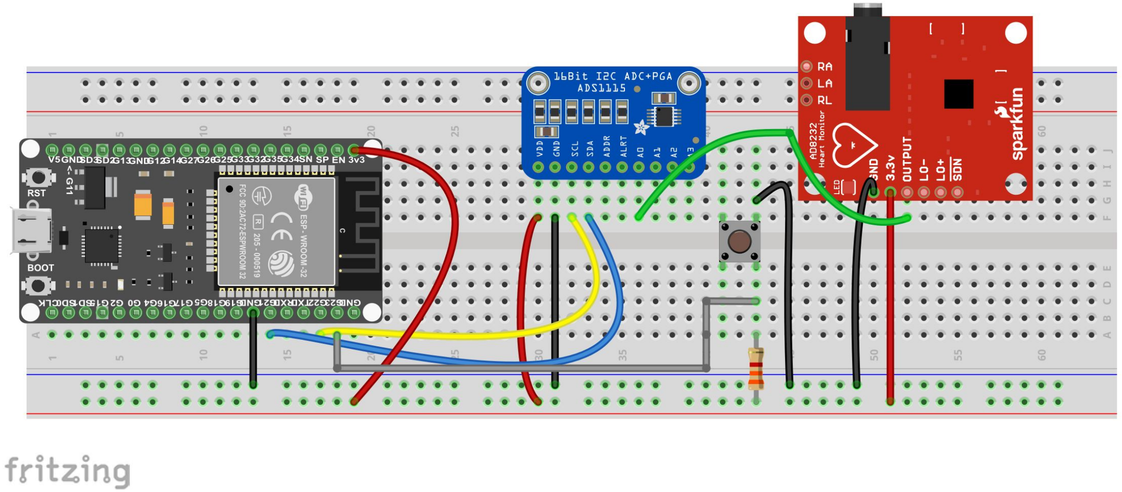


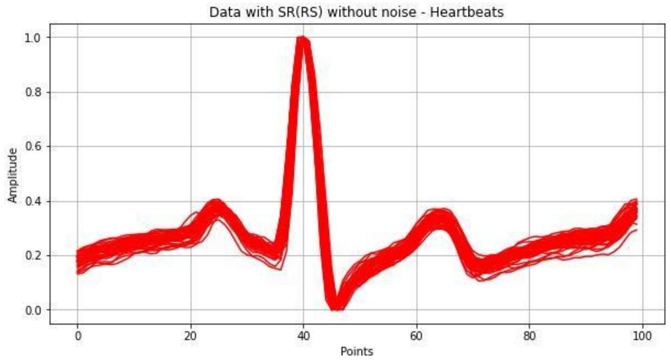
https://www.hackster.io/jallsonsuryo/livestock-wildlife-counting-from-drone-with-fomo-algorithm-a2f734

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](#)

Other TinyML / MCUs Project Examples

Vision

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

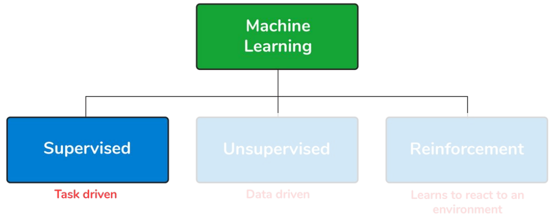
Sound

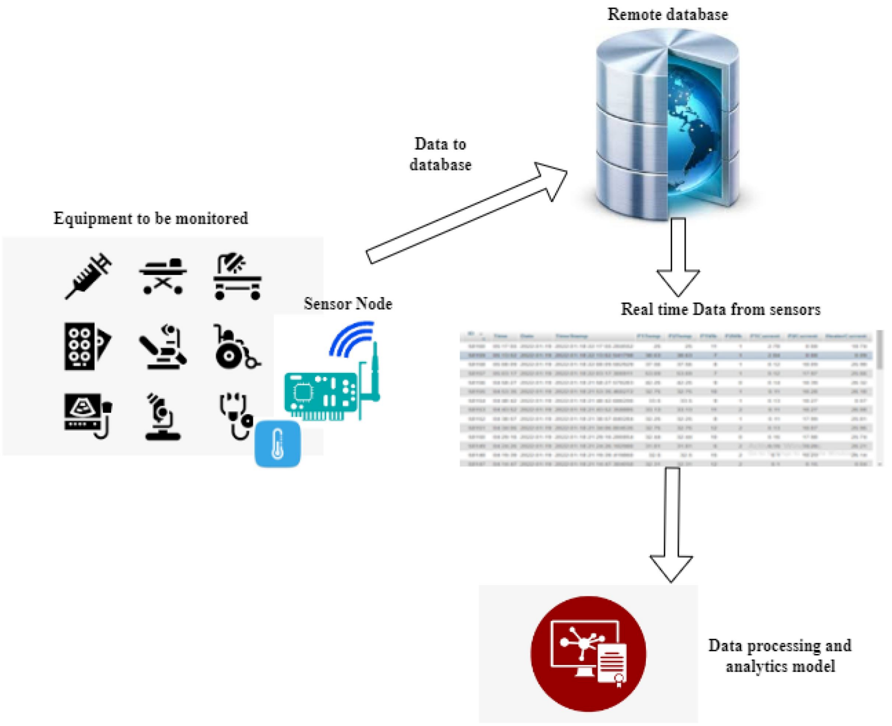
- Listening Temperature with Nano 33 [Doc]
- COPD Detection with Nano 33 [Doc]
- Sound Classification with XIAO BLE Sense [Doc]

Vibration

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

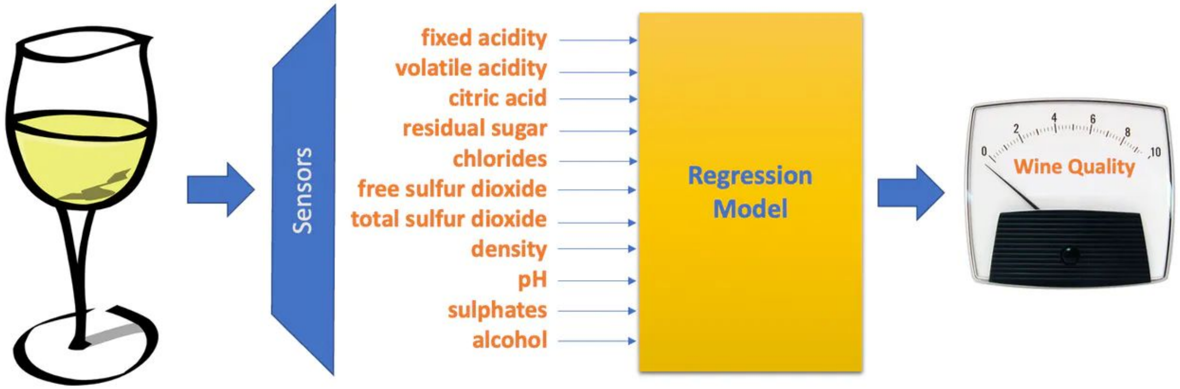
Regression on TinyML





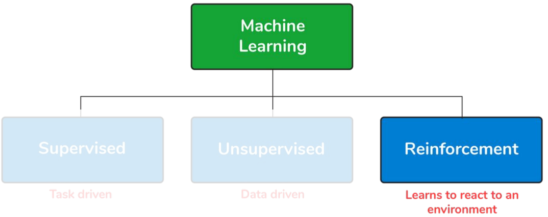
[On-Device IoT-Based Predictive Maintenance Analytics Model: Comparing TinyLSTM and TinyModel from Edge Impulse]()

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]

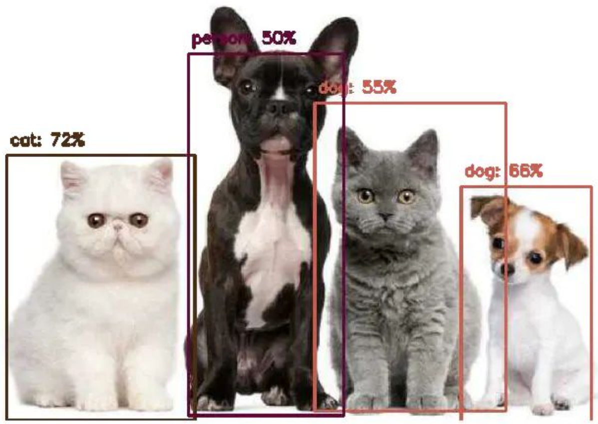
Aleksandra Faust[2] Guido C. H. E. de Croon[3] Vijay Janapa Reddi[1,4]

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

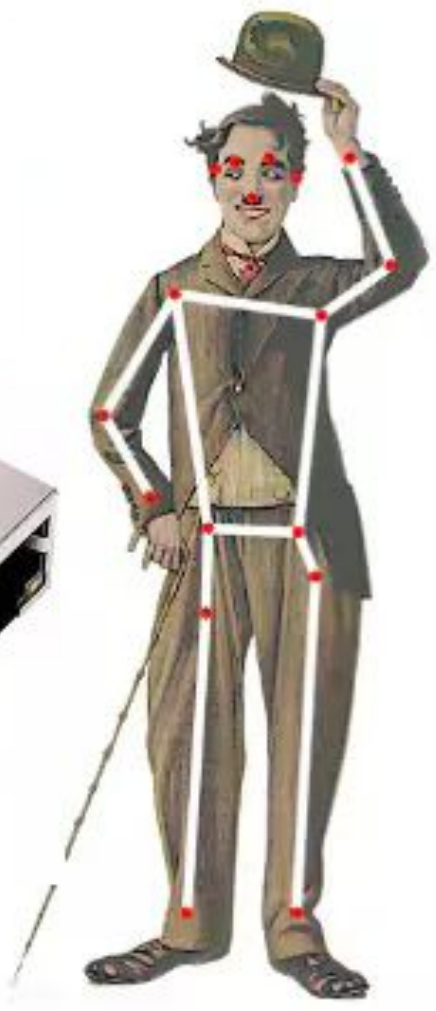
https://arxiv.org/abs/1909.11236

https://youtu.be/wmVKbX7MOnU

Exploring AI at the edge (Computer Vision)



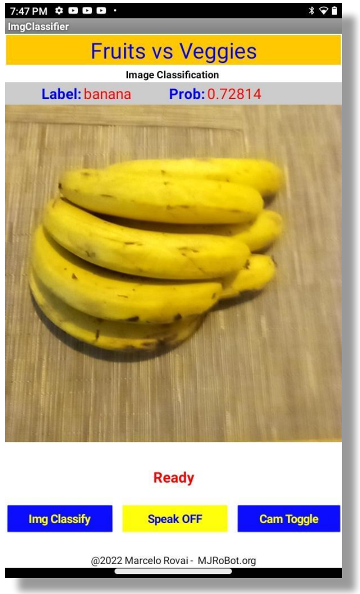
TensorFlow Lite

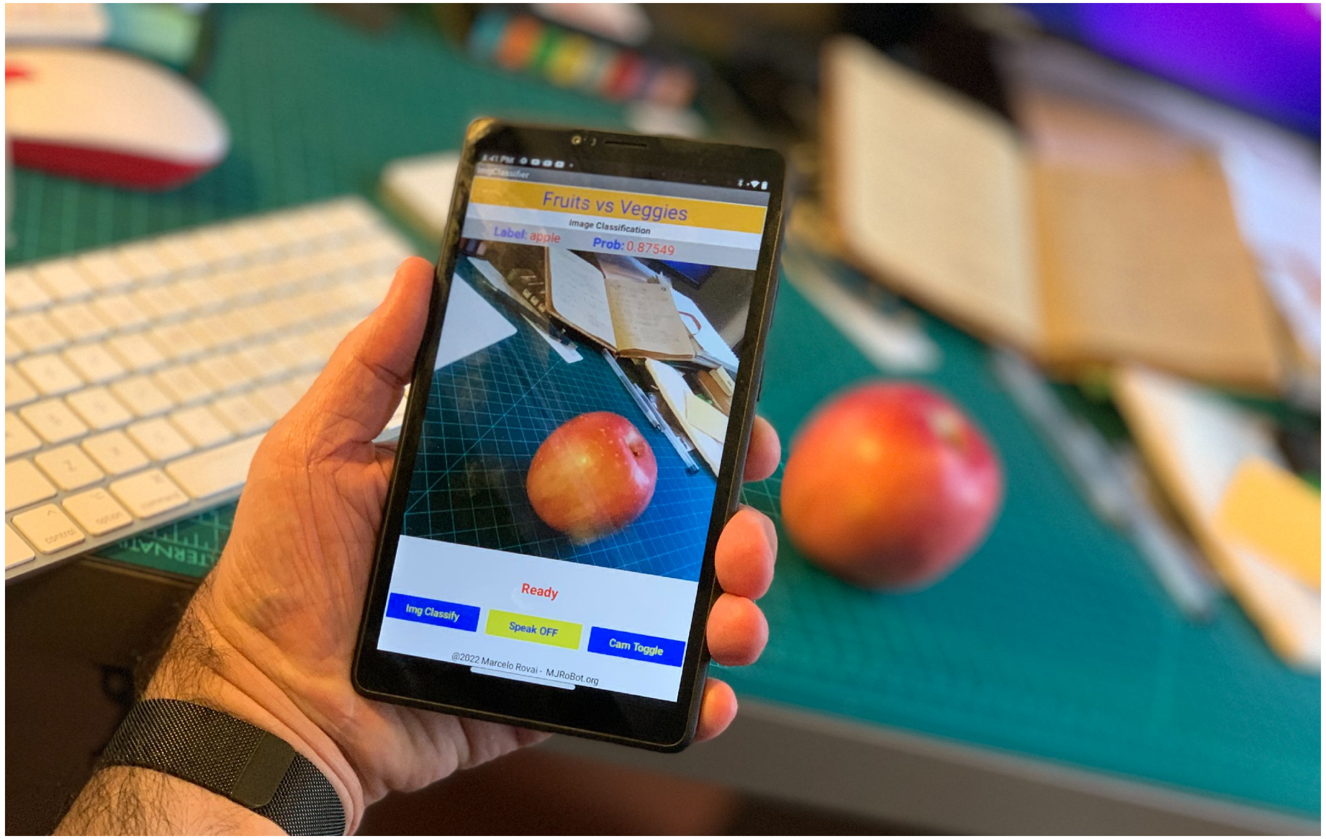


[Exploring AI at the Edge!](#)

Classifying Images using Smartphones

AI WITH APP INVENTOR

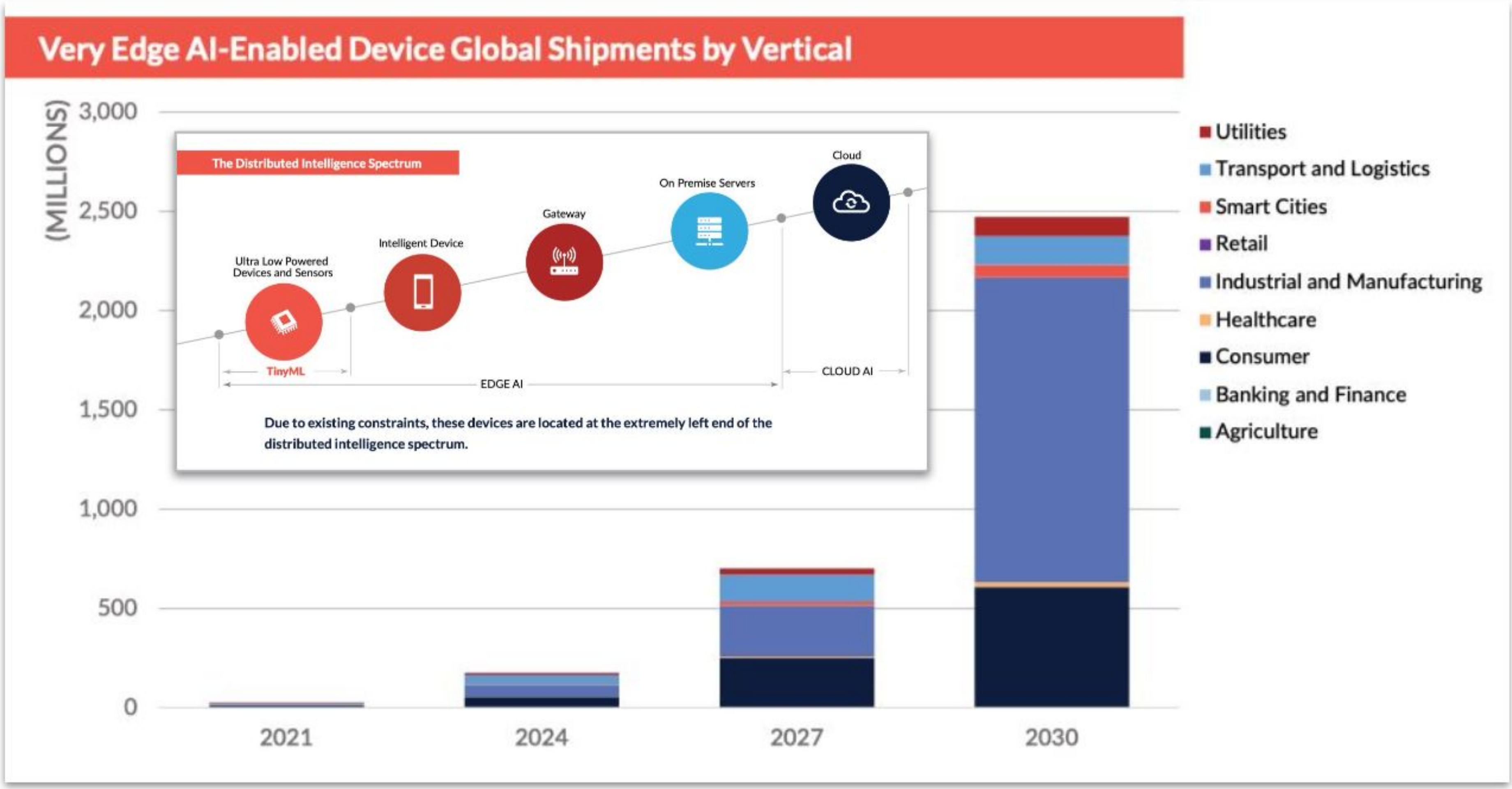




https://www.hackster.io/mjrobot/app-inventor-edgeml-image-classification-fruit-vs-veggies-b671da

The Future of the TinyML (Embedded ML)

Massive Potential for Impact



Source: ABI Research: TinyML

Conclusion

The Future of ML is Tiny and Bright

***Vijay Janapa Reddi**, Ph. D. | Associate Professor |*
John A. Paulson School of Engineering and Applied Sciences | Harvard University |

To learn more about Edge AI

- UNIFEI - IESTI01 TinyML - Machine Learning for Embedding Devices
- Professional Certificate in Tiny Machine Learning (TinyML) – edX/Harvard
- Introduction to Embedded Machine Learning - Coursera/Edge Impulse
- Computer Vision with Embedded Machine Learning - Coursera/Edge Impulse
- "Deep Learning with Python" book by François Chollet
- "TinyML" book by Pete Warden, Daniel Situnayake
- "TinyML Cookbook" by Gian Marco Iodice
- "AI at the Edge" book by Daniel Situnayake, Jenny Plunkett

Thanks

UNIFEI