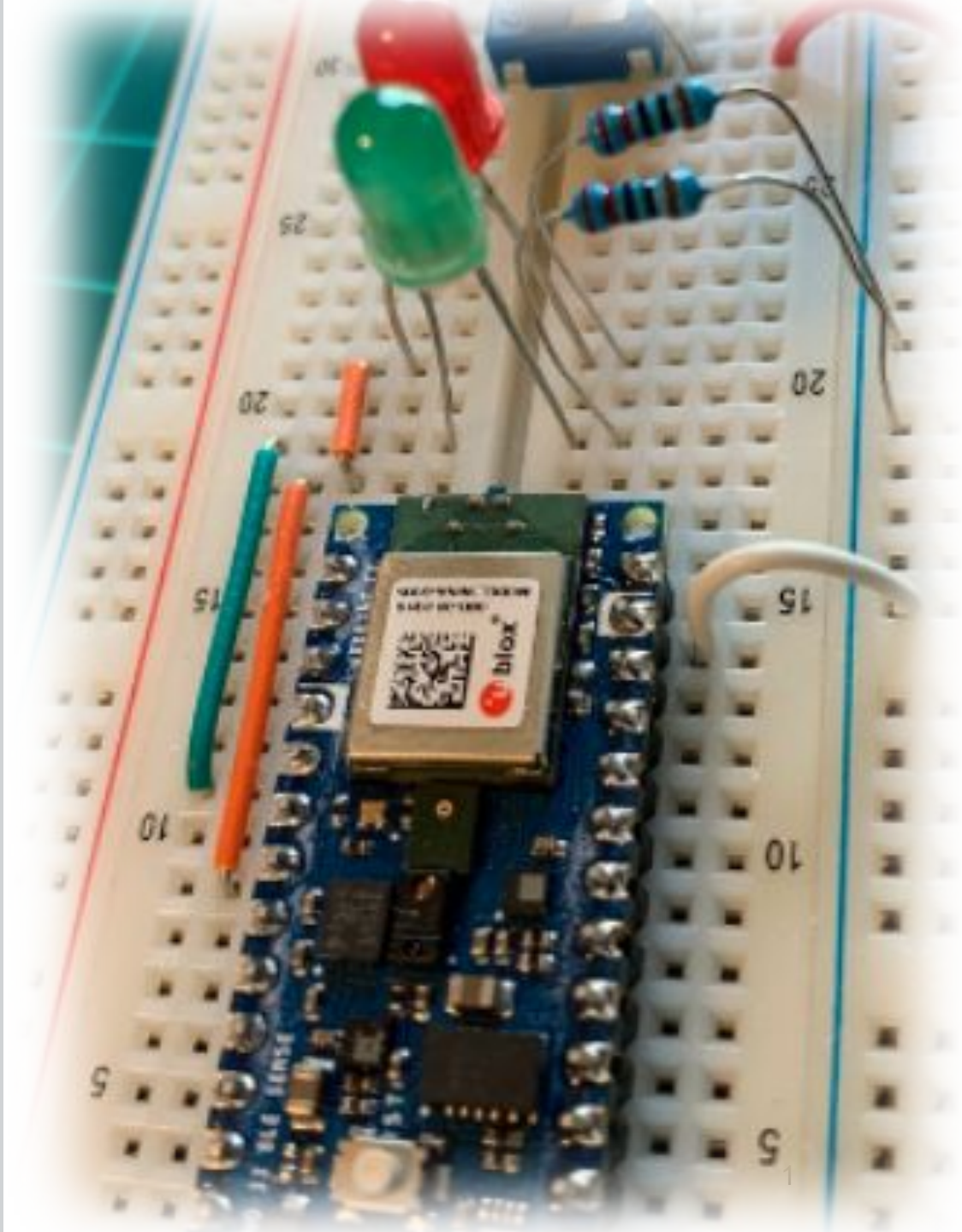
IESTI01 – TinyML

Embedded Machine Learning

1. About the Course & Syllabus



Prof. Marcelo Rovai
UNIFEI

Dear students,

Welcome to the first class of the IESTI01 (TinyML) EAD course. I am Professor Marcelo Rovai, a former student here at UNIFEI, and I am pleased to be with you this semester.

IESTI01 is a discipline that mixes Machine Learning (part of Artificial Intelligence) with small devices, such as microcontrollers and sensors, whose main characteristics are ultra-low power consumption, usually 32-bit CPUs, and a few kilobytes of memory.

We understand that the explosive growth of the Machine Learning field, the ease of use of software development platforms such as TensorFlow (TF), based on the Python language, and the current generation of powerful microcontrollers, make TinyML an indispensable topic of study for Engineering students in the areas of Electronics, as well as Computing and Control & Automation.

This mix of expertise and the pioneering nature of this discipline (We were the second university worldwide to have this type of course, Harvard School of Engineering being the first) leads us to significant challenges concerning the necessary basis for the minimum understanding of the matter. Thus, the time we have available for the course is short. So, We must commit ourselves to following the weekly classes and doing complementary activities such as readings, laboratories, and assignments. You can count on me to clarify doubts and to review the necessary concepts for a good understanding of the course (For that, use the FORUM, available in each class).

Greetings from the south of the world!

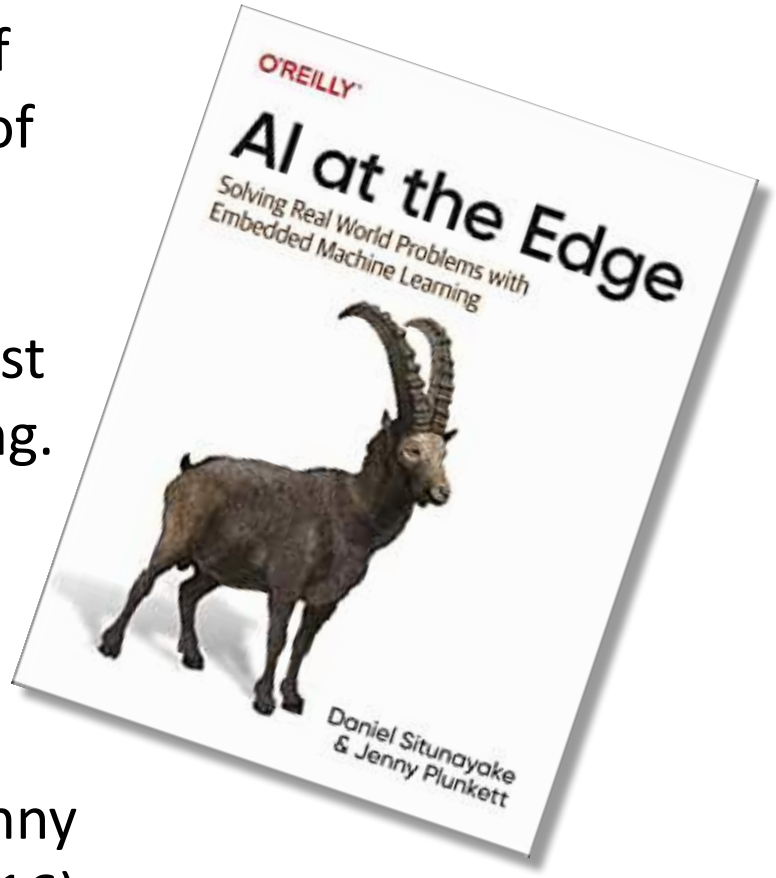
Prof. Marcelo Rovai

Who I am

- Brazilian from São Paulo, **Data Science Master's degree by UDD, Chile**, and MBA by IBMEC (INSPER), Brazil.
- Graduated in 1982 as an **Engineer from UNIFEI** with Specialization from Poli/USP, both in Brazil.
- Worked 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 Regional VP, and a Senior Advisor for Latin America.
- **Write about electronics**, publishing in sites as MJRoBot.org (Editor/Writer), Hackster.io (#1 Contributor), Instructables.com, and Medium.com (TDS – Towards Data Science).
- **Volunteer Professor** at UNIFEI Engineering Institute, teaching "Machine Learning applied to Embedded Devices" course (IESTI01).
- **TinyML4D group Co-Chair**, an initiative to bring TinyML education to developing countries.

Marcelo Rovai

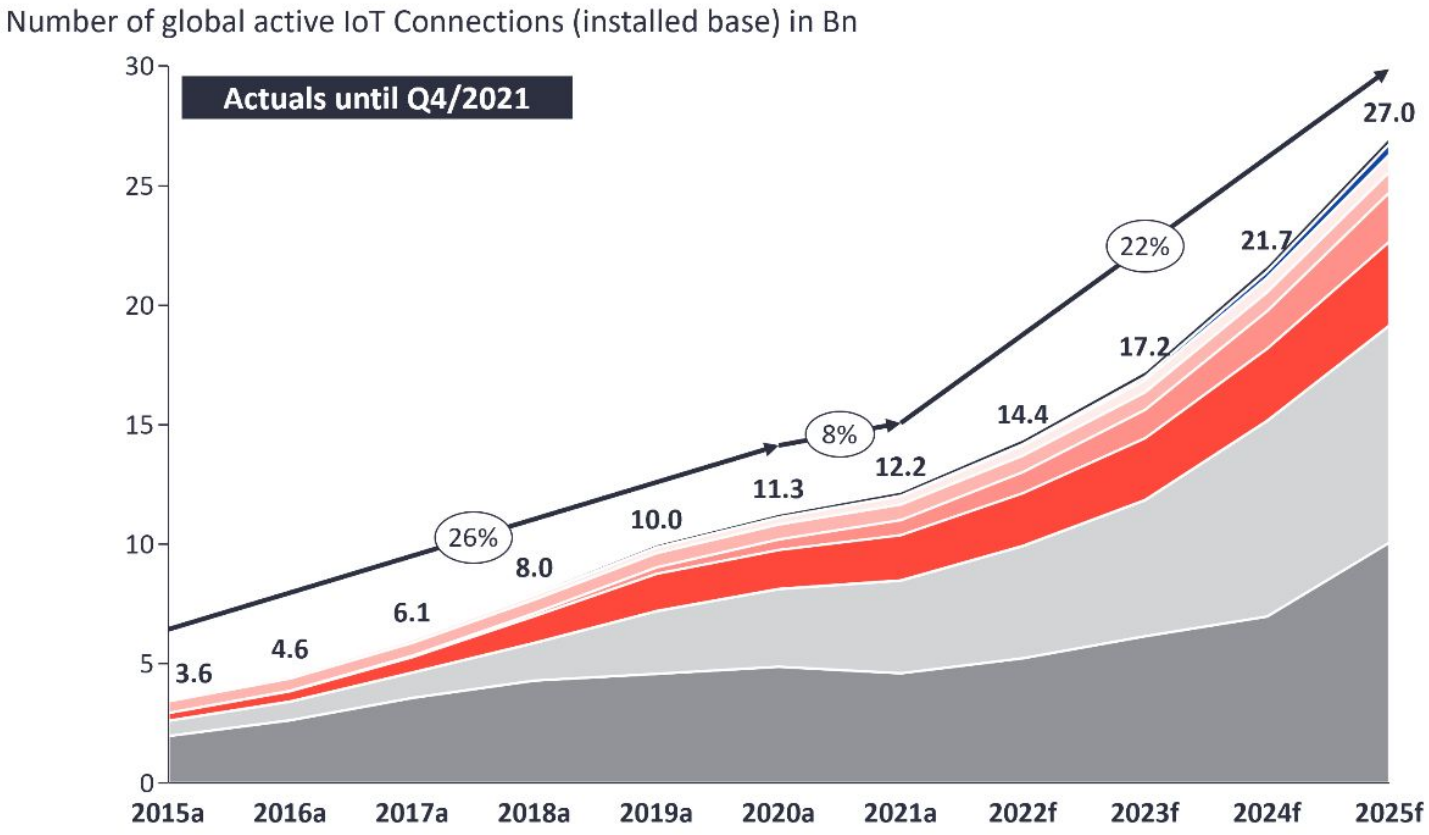
“**Edge AI** is a truly complete technology. As a topic, it makes use of knowledge from everything from the physical properties of semiconductor electronics all the way up to the engineering of high-level architectures that span devices and the cloud. It demands expertise in the most cutting-edge approaches to artificial intelligence and machine learning along with the most venerable skills of bare-metal embedded software engineering. It makes use of the entire history of computer science and electrical engineering, laid out end to end.”



Situnayake, Daniel; Plunkett, Jenny
AI at the Edge (pp. 215-216)
O'Reilly Media

Internet of Things (IoT)

Global IoT Market Forecast [in billion connected IoT devices]



CONNECTIVITY TYPE	CAGR 20-21	CAGR 21-25
Wireless Neighborhood Area Networks (WNAN)	17%	11%
5G IoT	-	159%
Other	22%	20%
Wired IoT	4%	7%
LPWA	42%	34%
Legacy Cellular (2G/3G/4G)	16%	17%
Wireless Local Area Networks (WLAN)	19%	24%
Wireless Personal Area Networks (WPAN)	-6%	22%

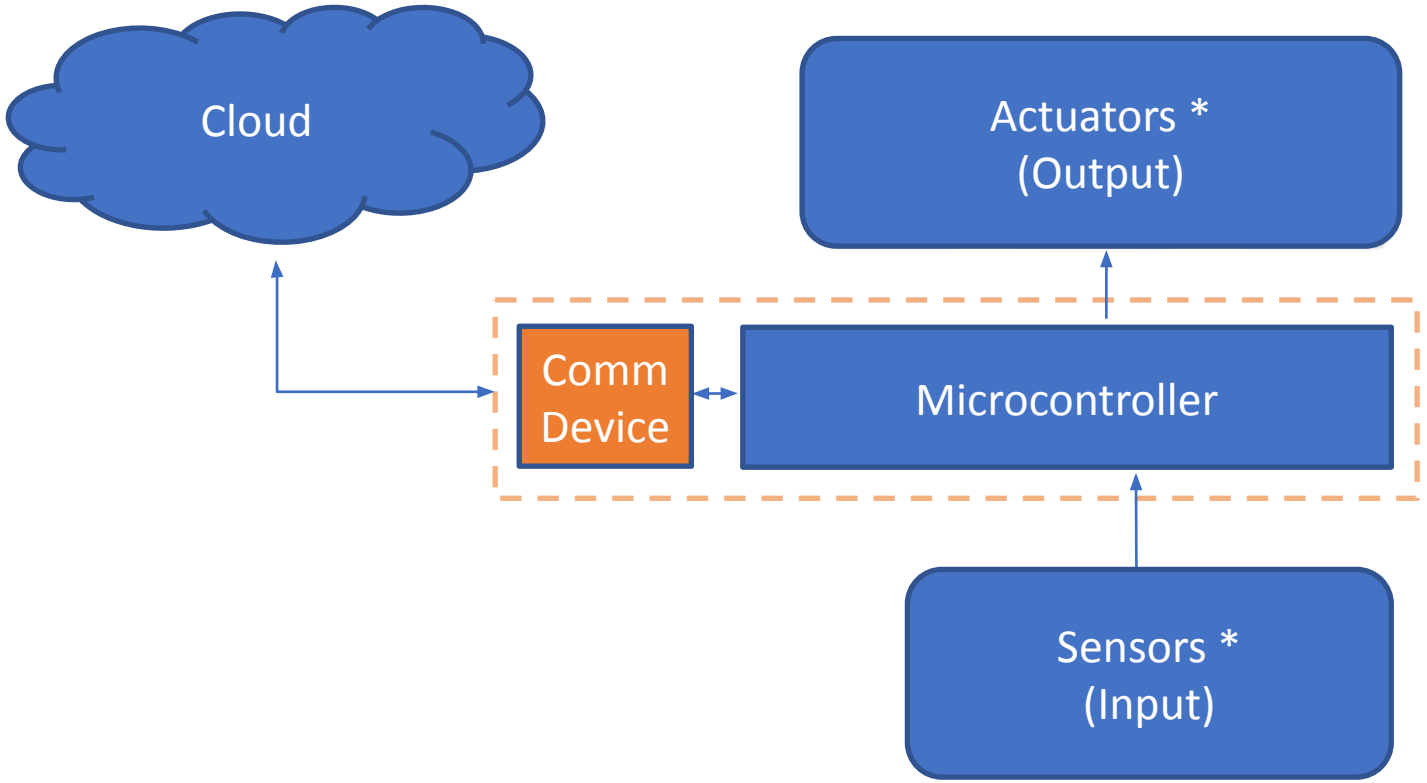
XX% = CAGR

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.

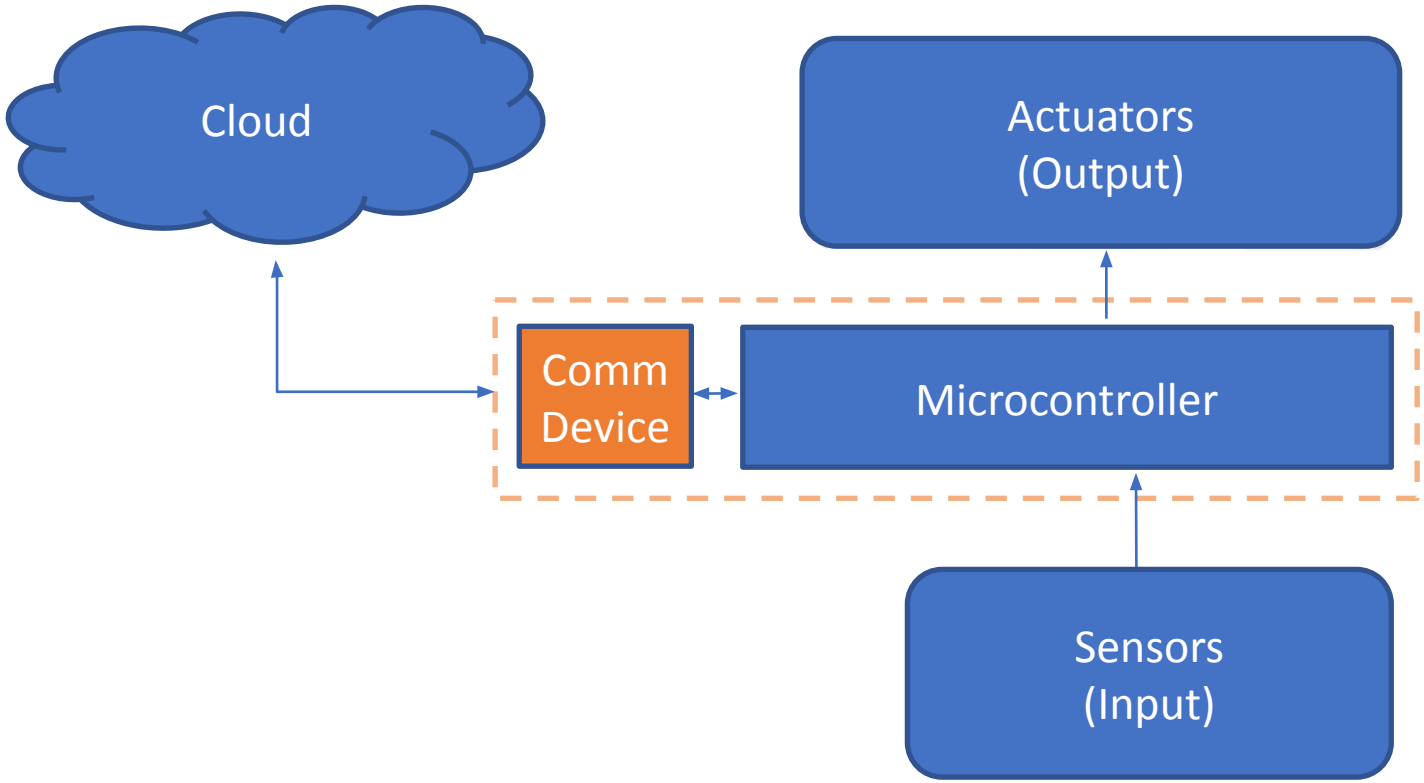
https://iot-analytics.com/number-connected-iot-devices

Typical IoT Project



* “Things”

Typical IoT Project



5 Quintillion

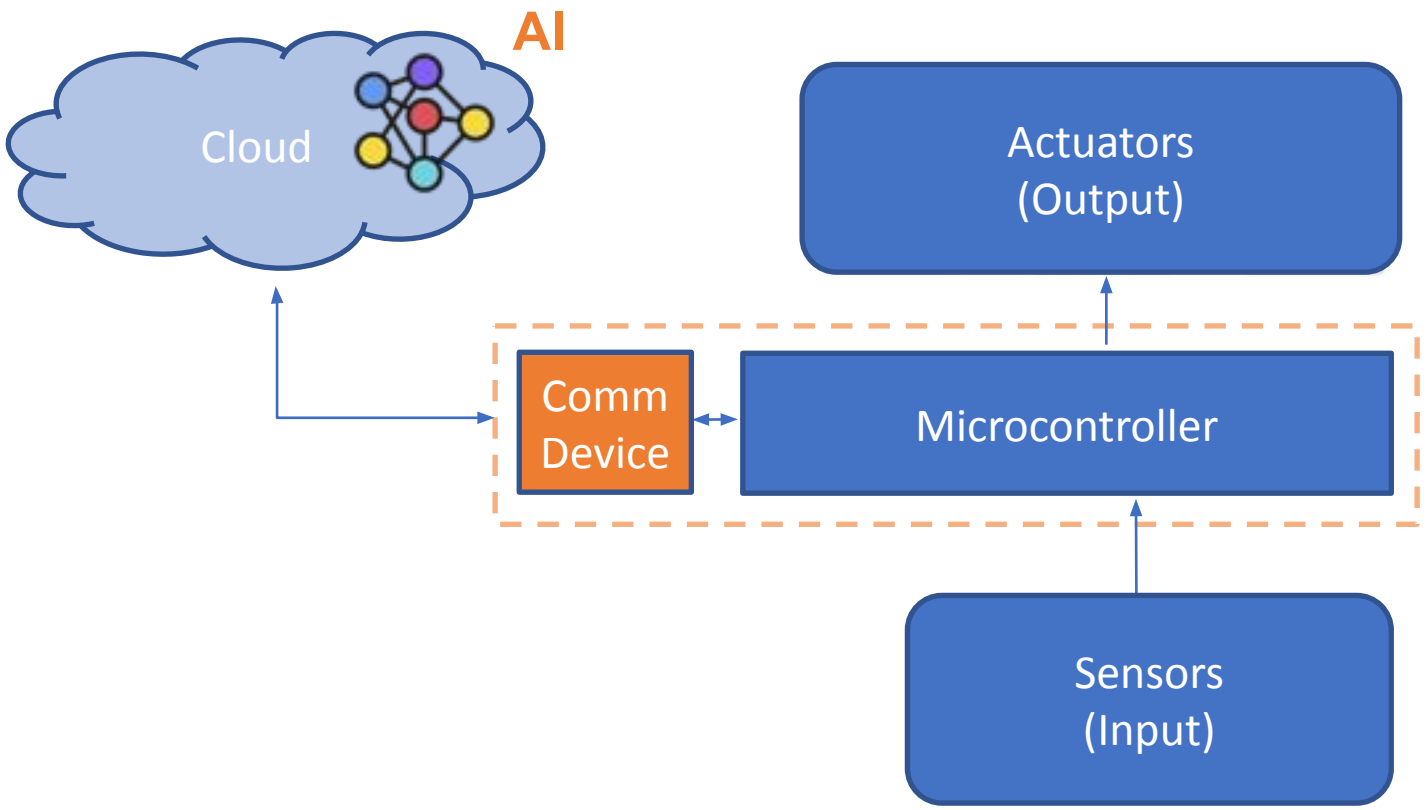
bytes of data produced every day by IoT

<1%

of unstructured data is analyzed or used at all

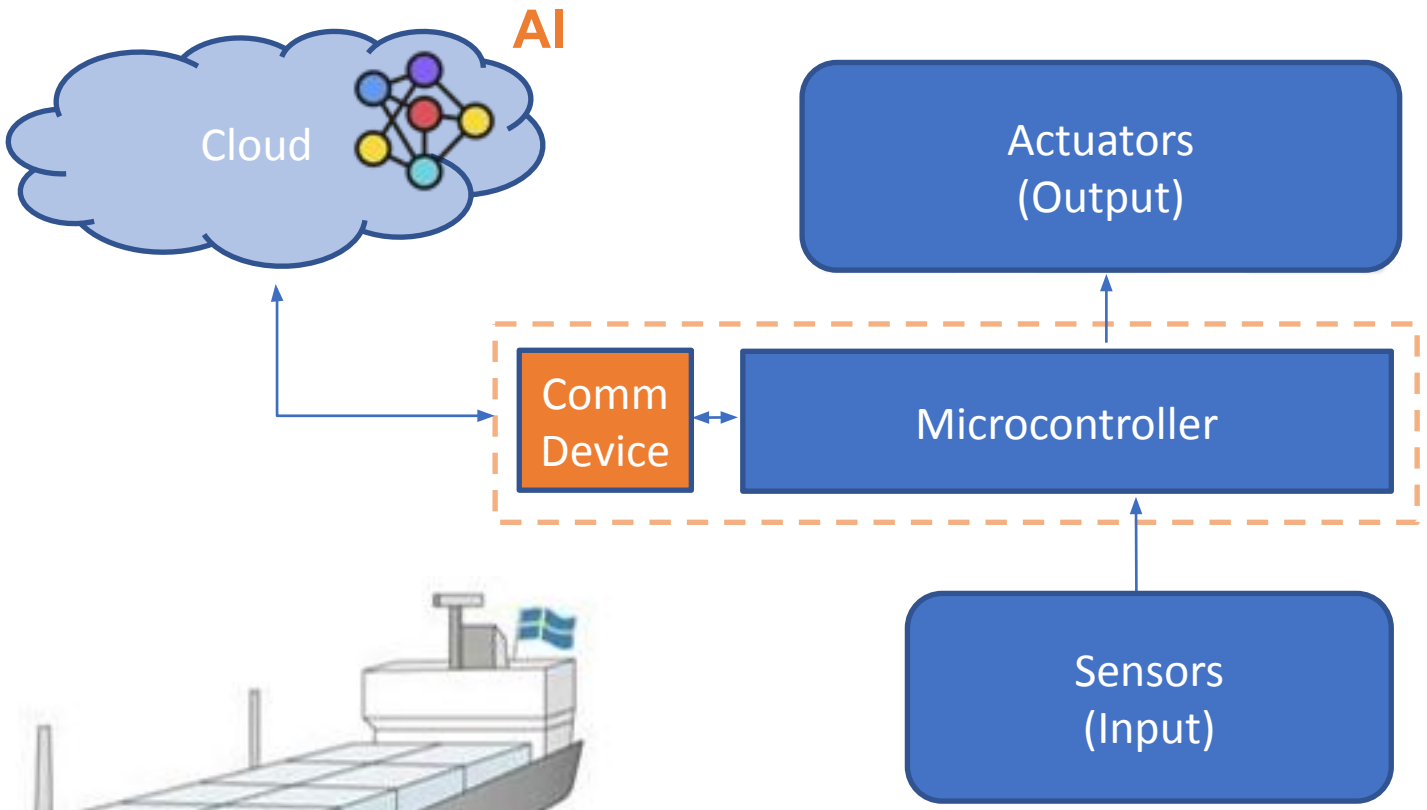
Source: Harvard Business Review, What's Your Data Strategy?, April 18, 2017
Cisco, Internet of Things (IoT) Data Continues to Explode Exponentially. Who Is Using That Data and How?, Feb 5, 2018

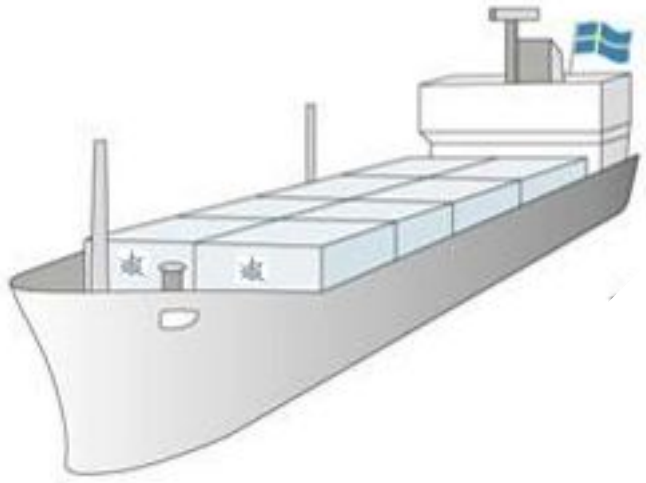
Typical AIoT Project



Typical AIoT Project …

… Issues

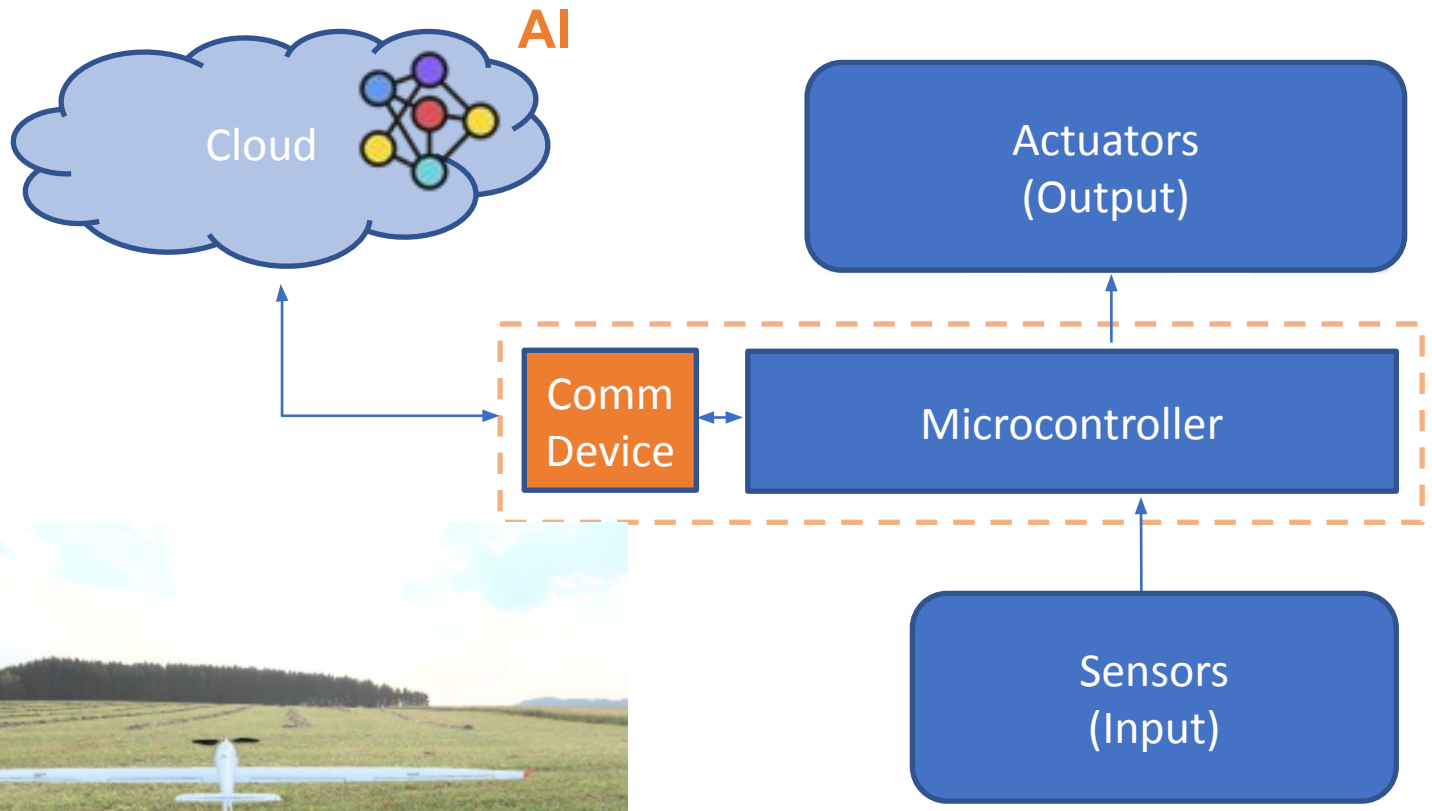




Bandwidth

Typical AIoT Project …

… Issues

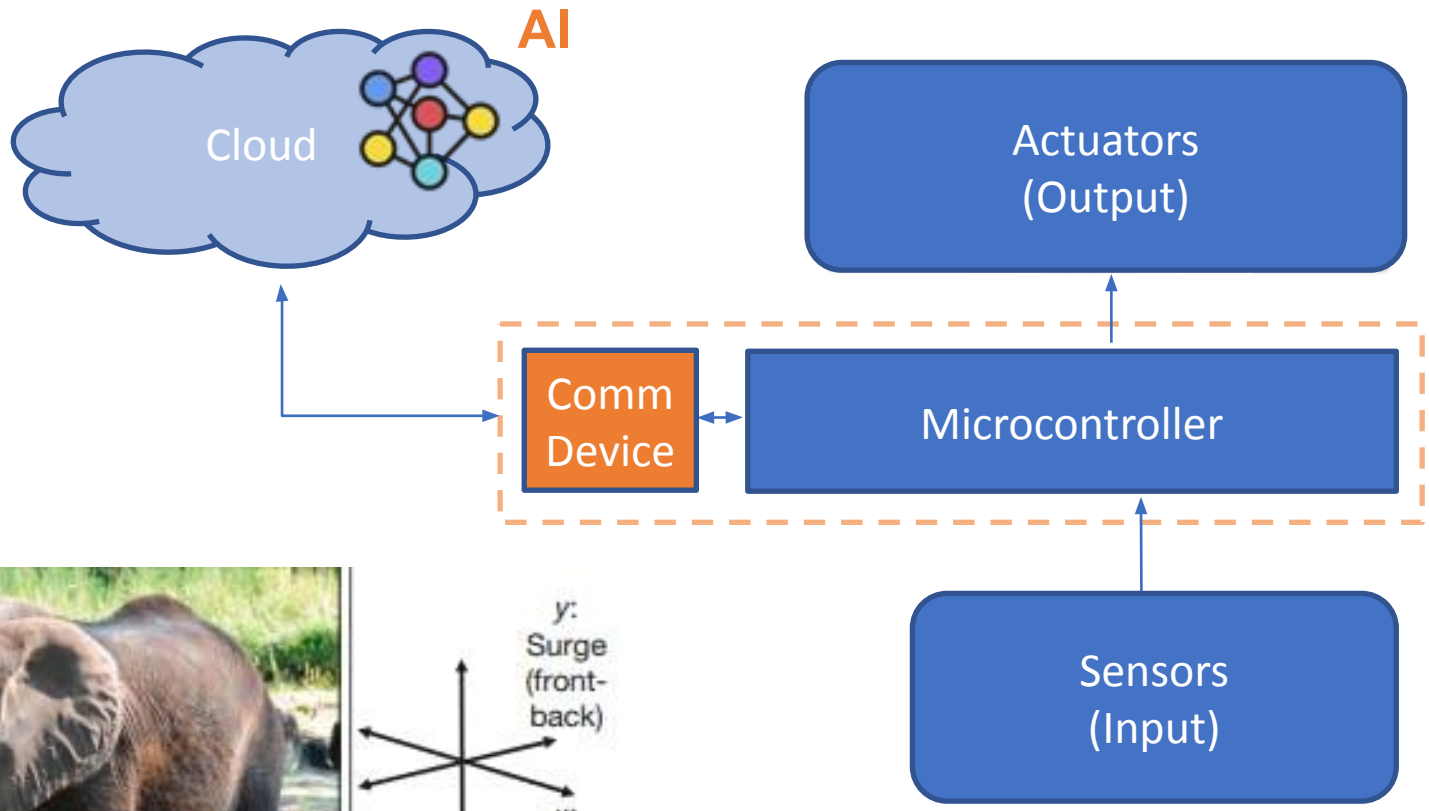


Bandwidth

Latency

Typical AIoT Project …

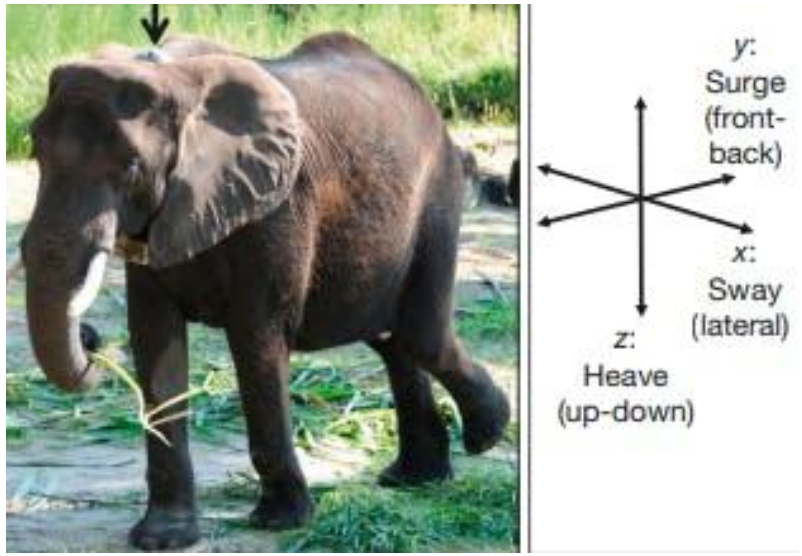
… Issues



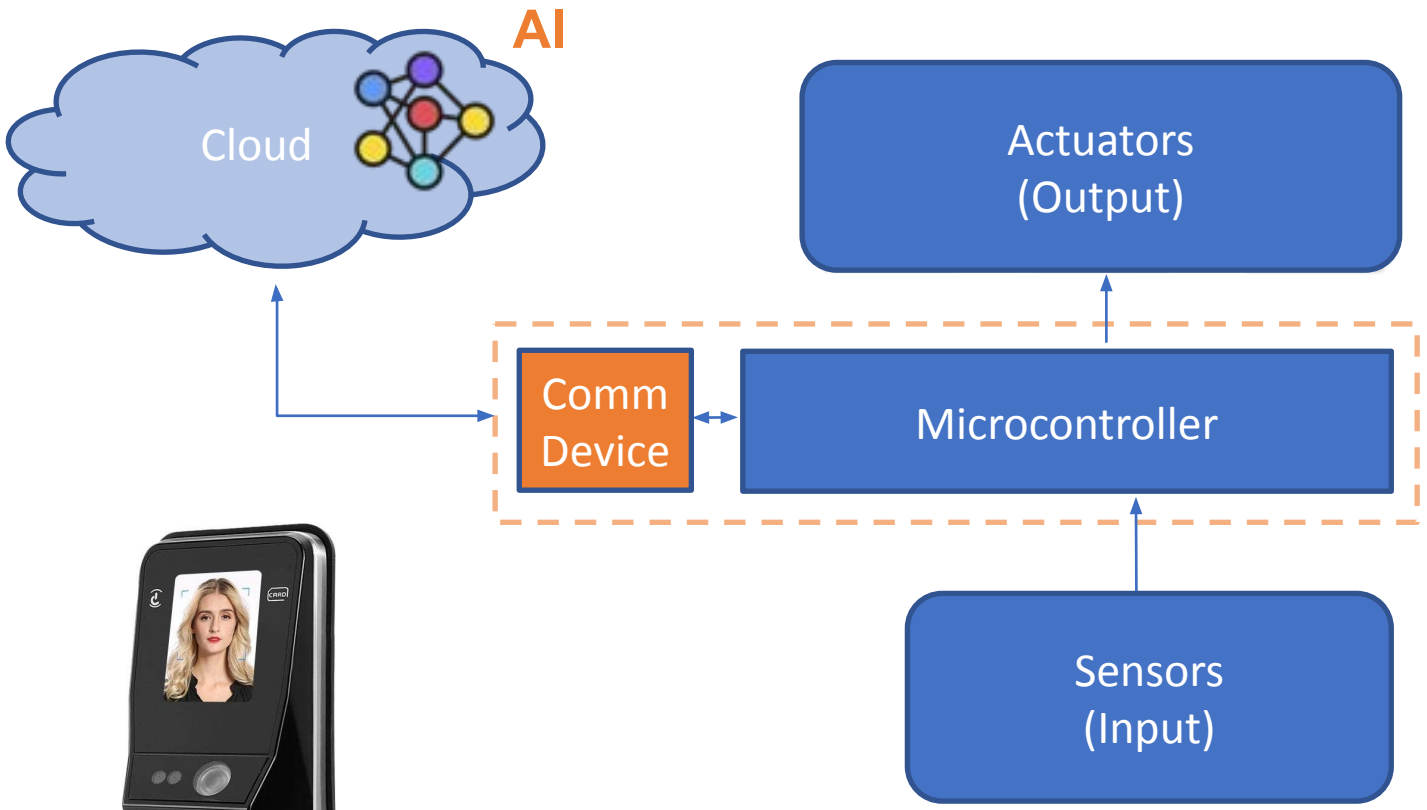
Bandwidth

Latency

Energy



Typical AIoT Project …



… Issues

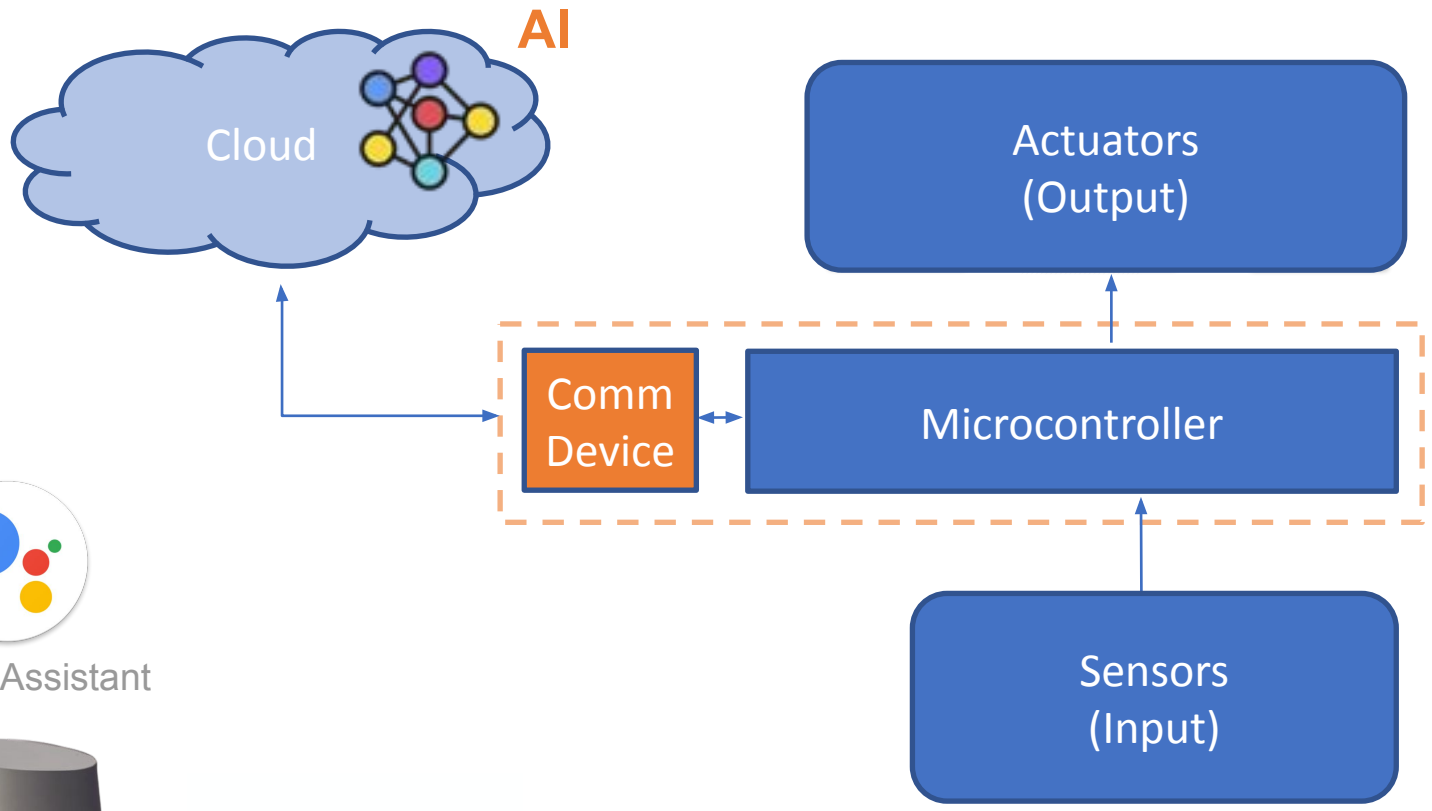
Bandwidth

Latency

Energy

Reliability

Typical AIoT Project …



Google Assistant

… Issues

Bandwidth

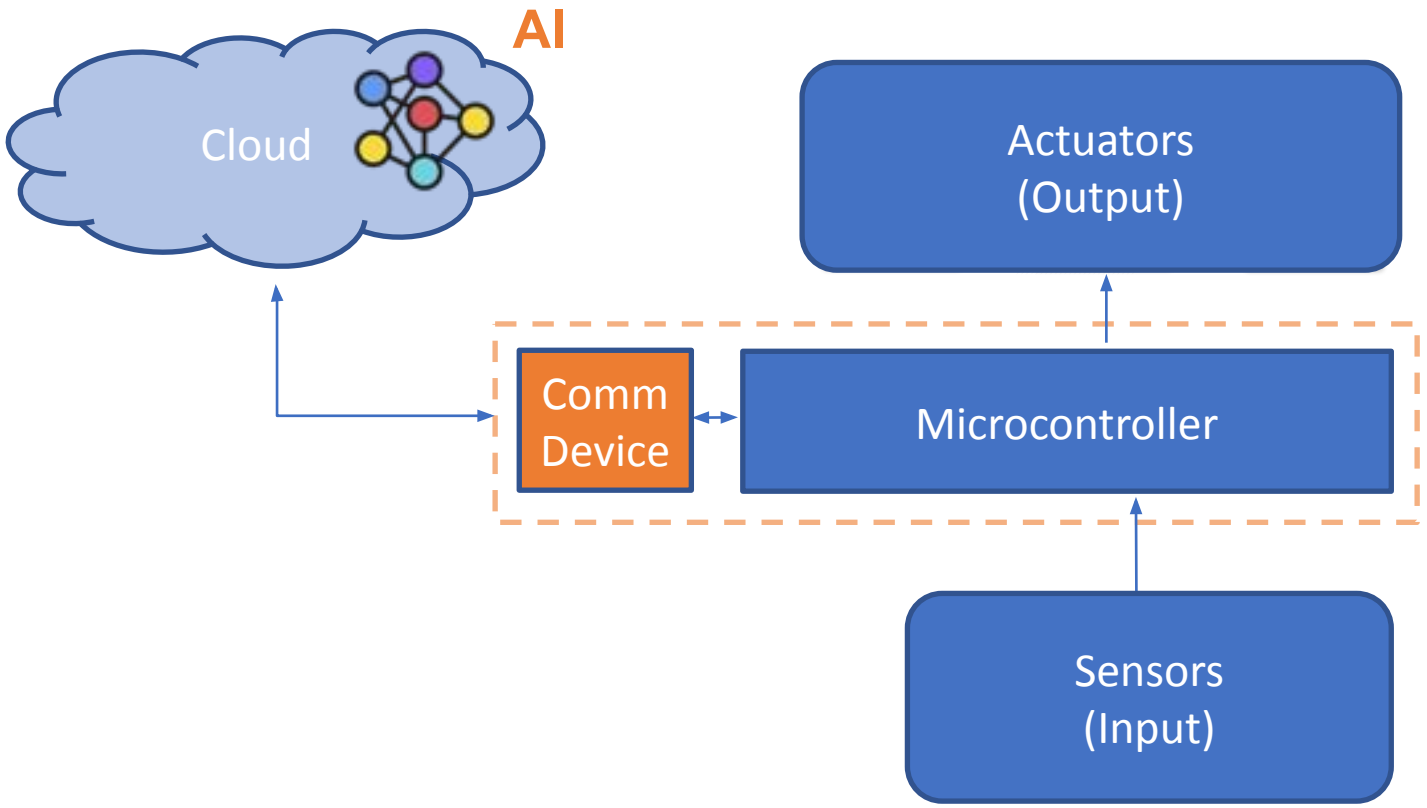
Latency

Energy

Reliability

Privacy

Typical AIoT Project …



… Issues

Bandwidth

Latency

Energy

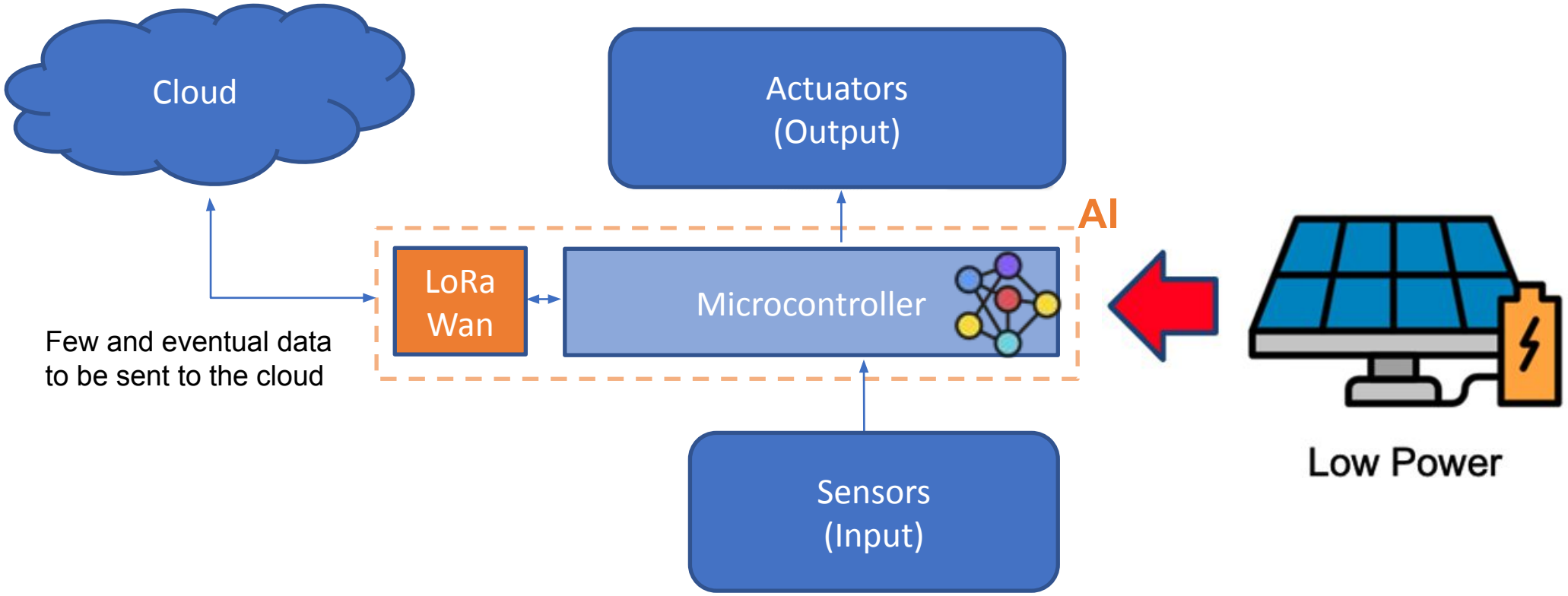
Reliability

Privacy

… Solution ?

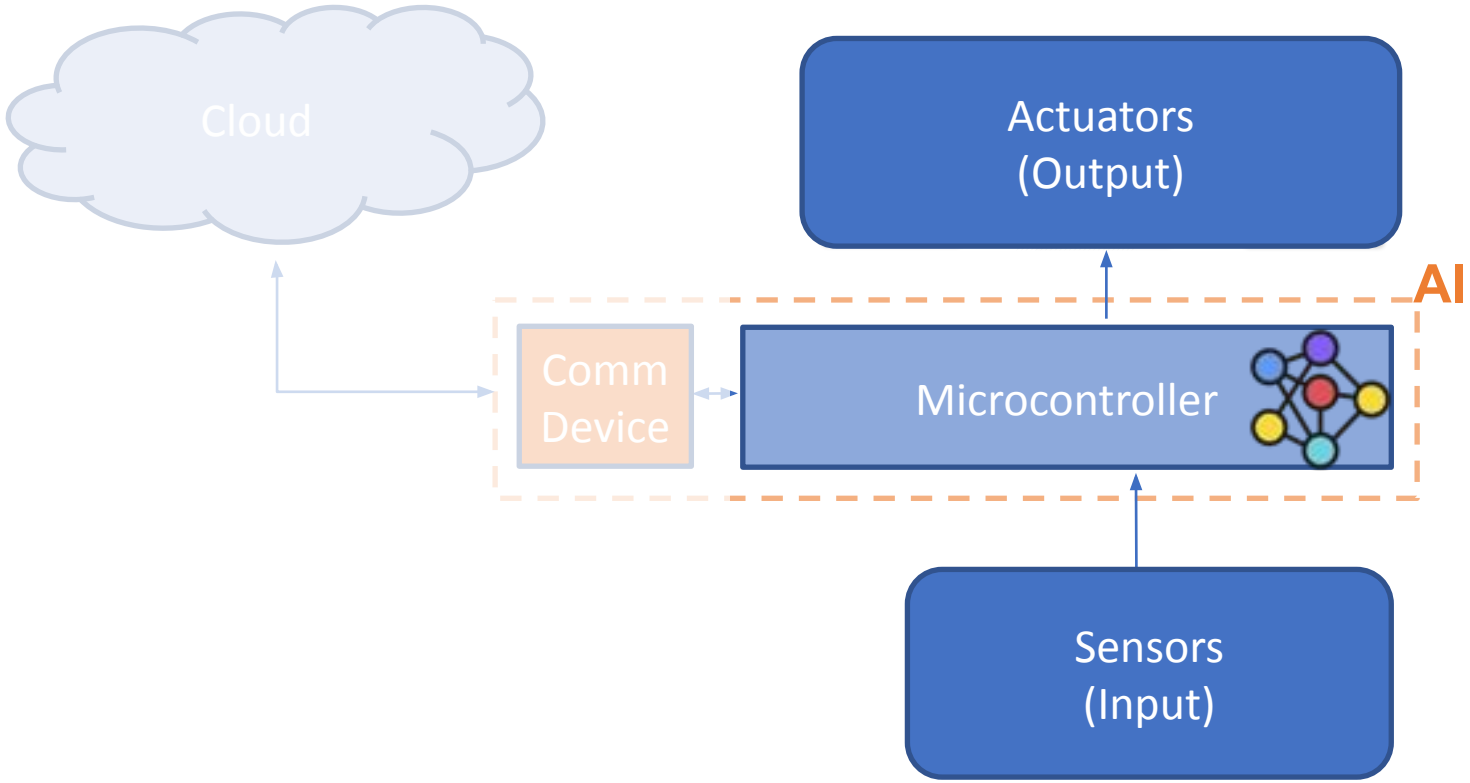
IoT 2.0 * – Edge AI/ML

* Intelligence of Things



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

When to use an Edge AI/ML approach:



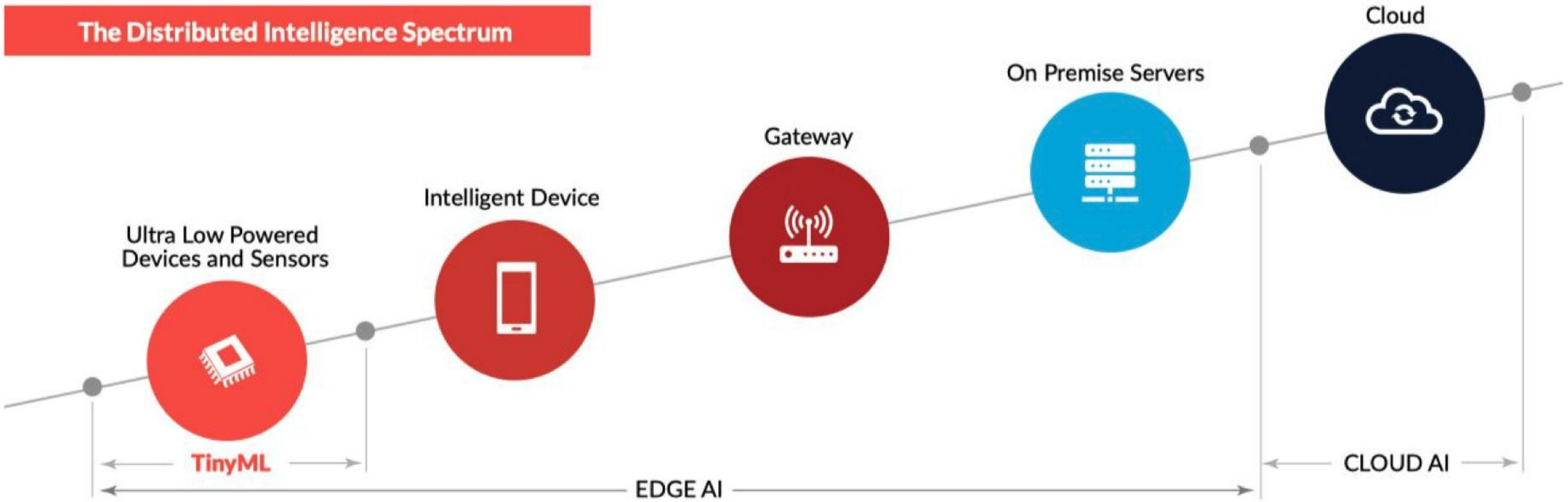
Bandwidth

Latency

Energy

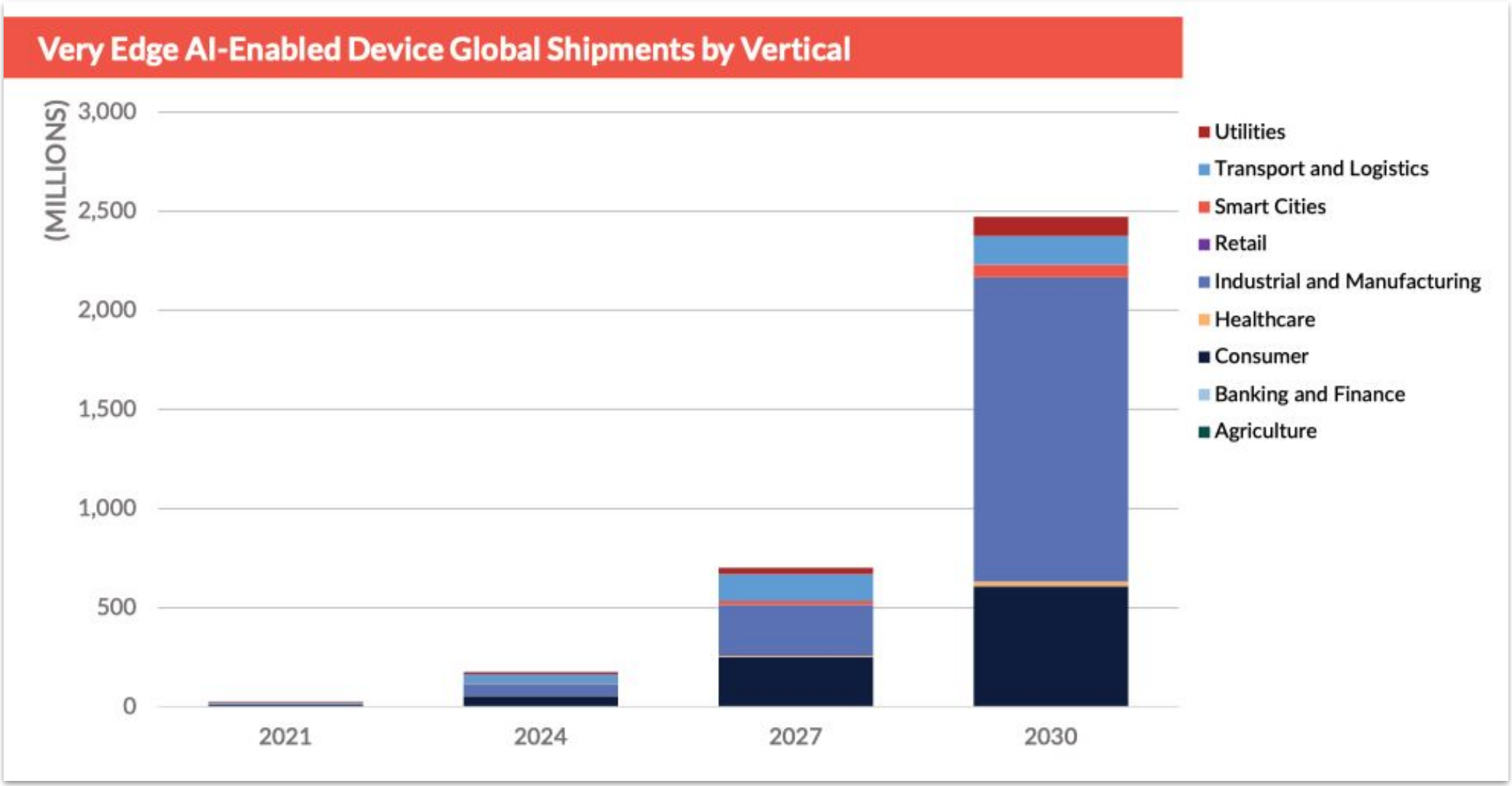
Reliability

Privacy



Source: ABI Research: TinyML

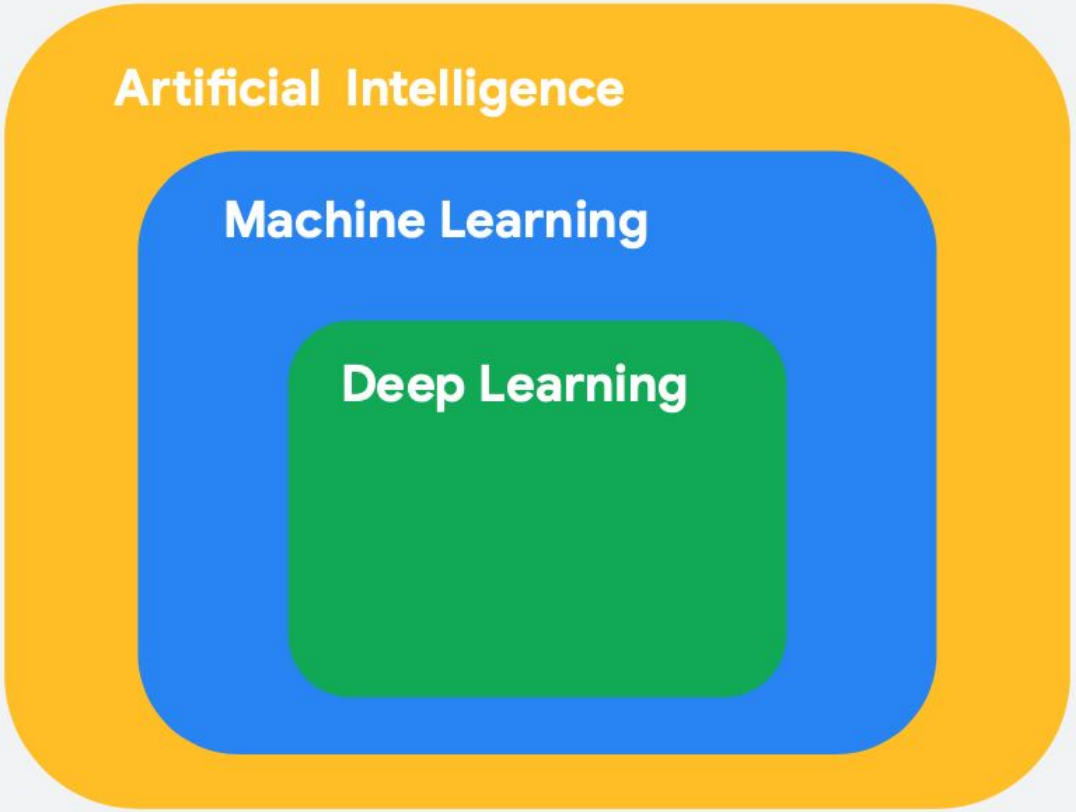
Market Forecast



Source: ABI Research: TinyML

Embedded ML (TinyML)

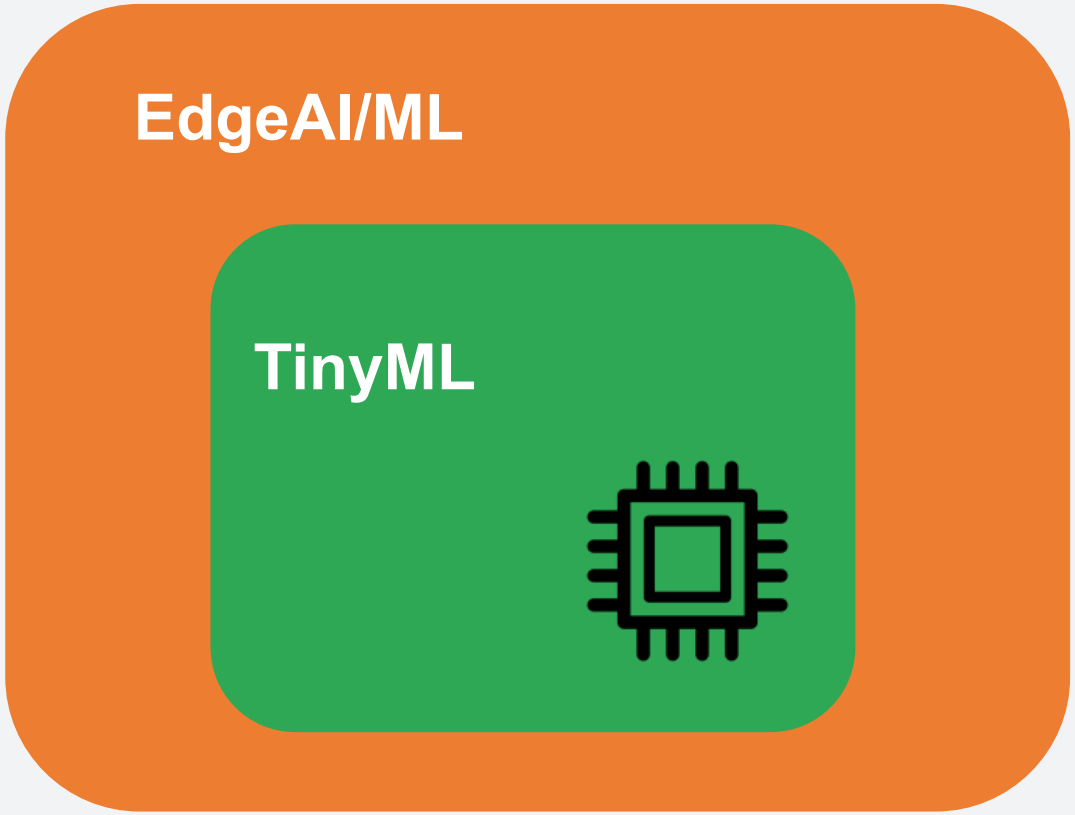
Introduction



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**

Algorithms, hardware, software

What is Tiny Machine Learning (TinyML)?

TinyML

Fastest-growing field of **ML**

Algorithms, hardware, software

On-device sensor analytics

What is Tiny Machine Learning (TinyML)?

TinyML

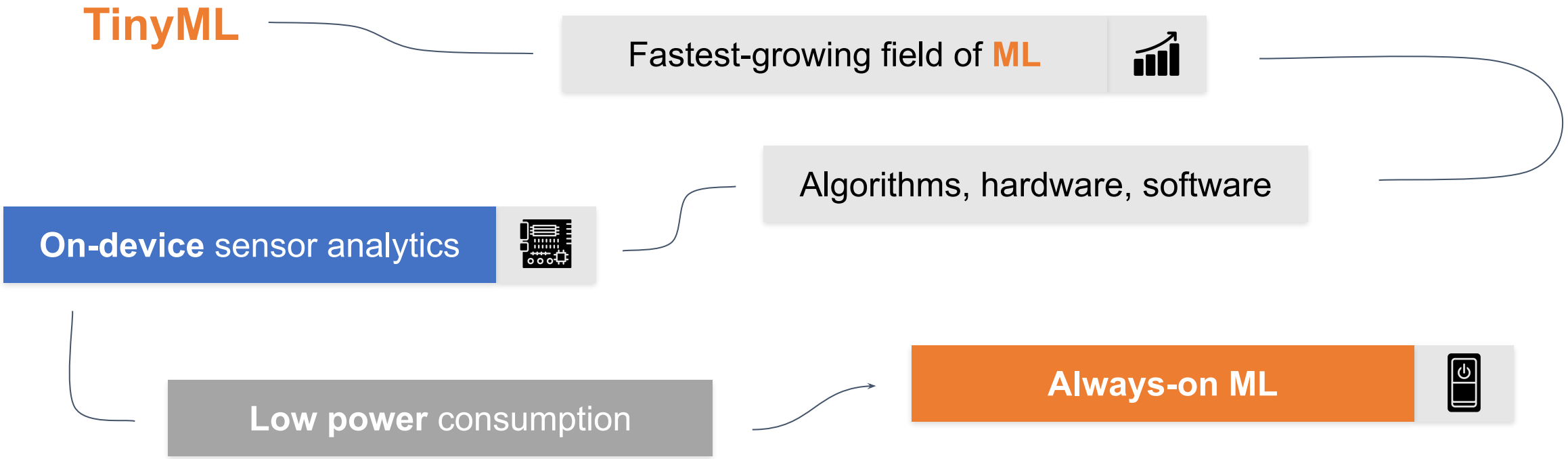
Fastest-growing field of **ML**

Algorithms, hardware, software

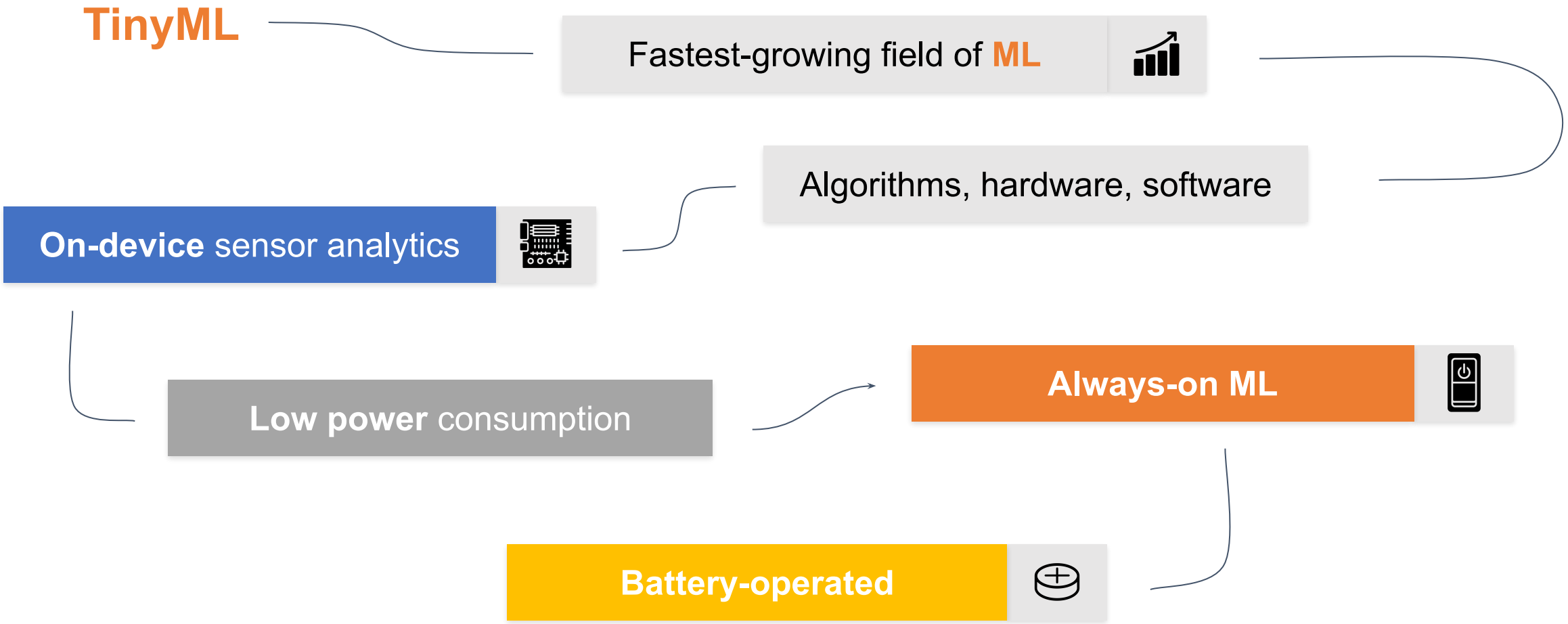
On-device sensor analytics

Low power consumption

What is Tiny Machine Learning (TinyML)?



What is Tiny Machine Learning (TinyML)?



Hardware

Power

EdgeML

TinyML

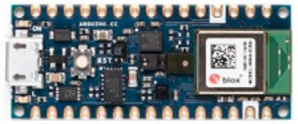
Video Classification 2 MB+

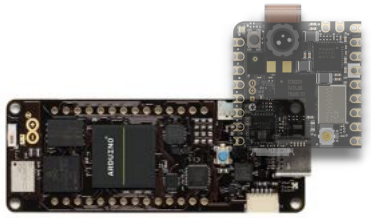
Object Detection Complex Voice Processing 1 MB+

Image Classification 250 KB+

KeyWord Spotting Audio Classification 50 KB

Anomaly Detection Sensor Classification 20 KB





Rpi-Pico (Cortex-M0+)

Arduino Nano (Cortex-M4)

Arduino Pro (Cortex-M7)

RaspberryPi SmartPhone (Cortex-A)

Jetson Nano (Cortex-A + GPU)

Source: Edge Impulse

Hardware

Power

EdgeML

TinyML

Video Classification 2 MB+

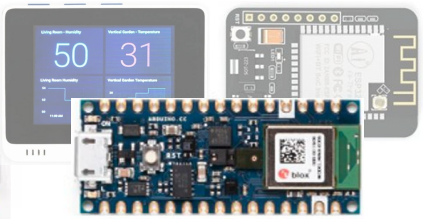
Object Detection Complex Voice Processing 1 MB+

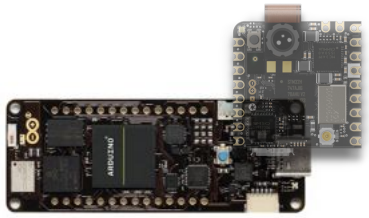
Image Classification 250 KB+

KeyWord Spotting Audio Classification 50 KB

Anomaly Detection Sensor Classification 20 KB

Wio ESP32





Rpi-Pico (Cortex-M0+)

Arduino Nano (Cortex-M4)

Arduino Pro (Cortex-M7)

RaspberryPi SmartPhone (Cortex-A)

Jetson Nano (Cortex-A + GPU)

Source: Edge Impulse

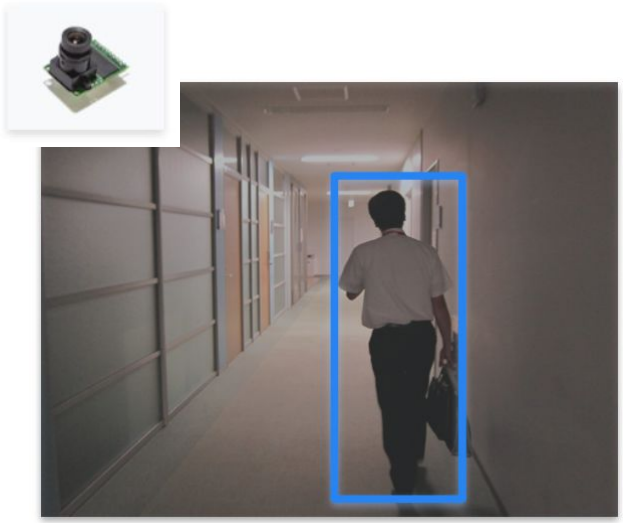
TinyML Application

Examples

Sound

Vibration

Vision

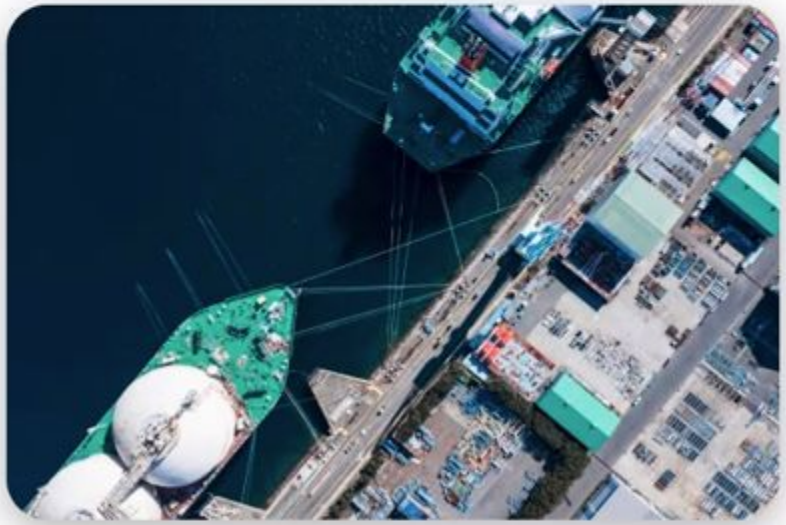


Predictive Maintenance

Motion, current, audio and camera

- → **Industrial**
- → White goods
- → Infrastructure
- → Automotive

Asset Tracking & Monitoring



Motion, temp, humidity, position, audio and camera

- → Logistics
- → Infrastructure
- → Buildings
- → Agriculture

Human & Animal Sensing

Motion, radar, audio, PPG, ECG

- → Health
- → Consumer
- → Industrial

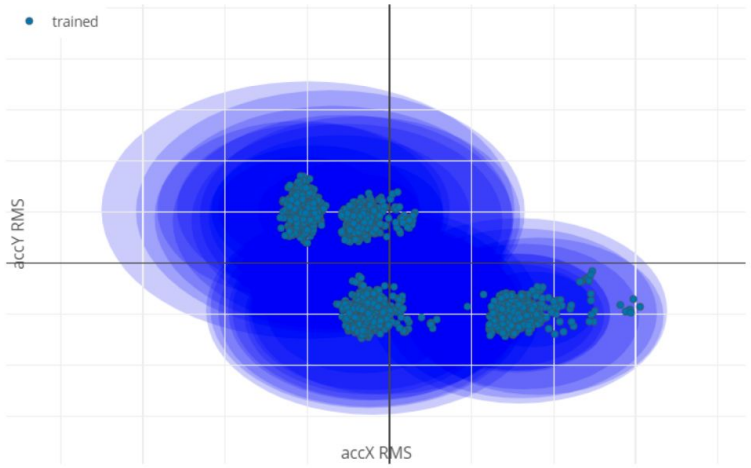
EDGE IMPULSE

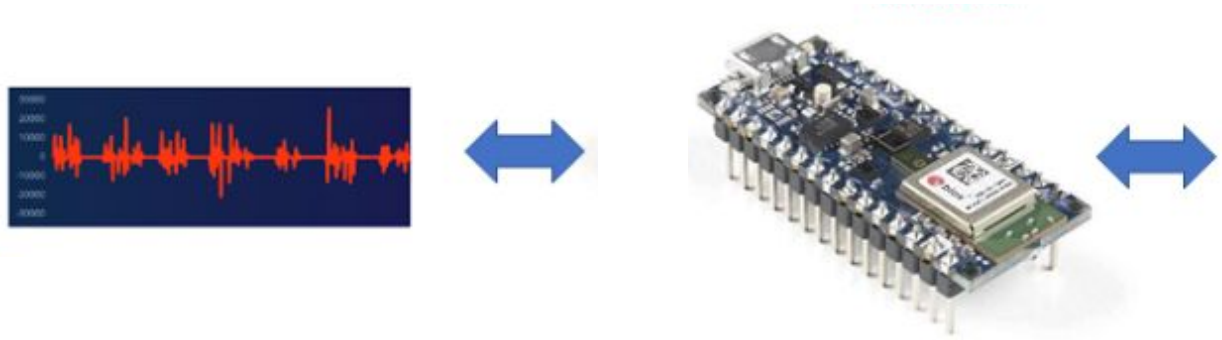
Application: Factory machinery

Ball Bearings

Accelerometer

Industry – Anomaly Detection





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

Cow Monitoring

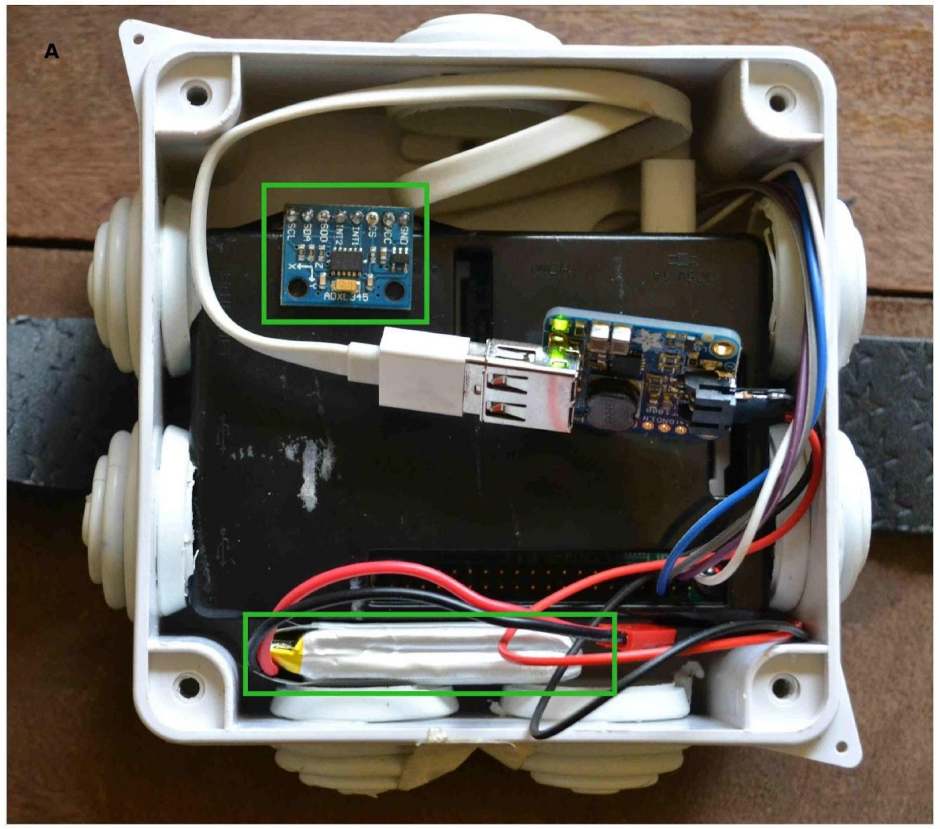
Using the Internet of Things for Agricultural Monitoring

"We aim to deploy a variety of sensors for agricultural monitoring. One of the projects involves using accelerometer sensors to monitor activity levels in dairy cows with a view to determining when the cows are on heat or when they are sick."

Ciira wa Maina, Ph.D.

Senior Lecturer
Department of Electrical and Electronic Engineering
Dedan Kimathi University of Technology
Nyeri Kenya
Email: ciira.maina@dkut.ac.ke

Kenia



https://sites.google.com/site/cwamainadekut/research

Coffee Disease Classification

João Vitor Yukio Bordin Yamashita
Graduando em Engenharia Eletrônica pela UNIFEI

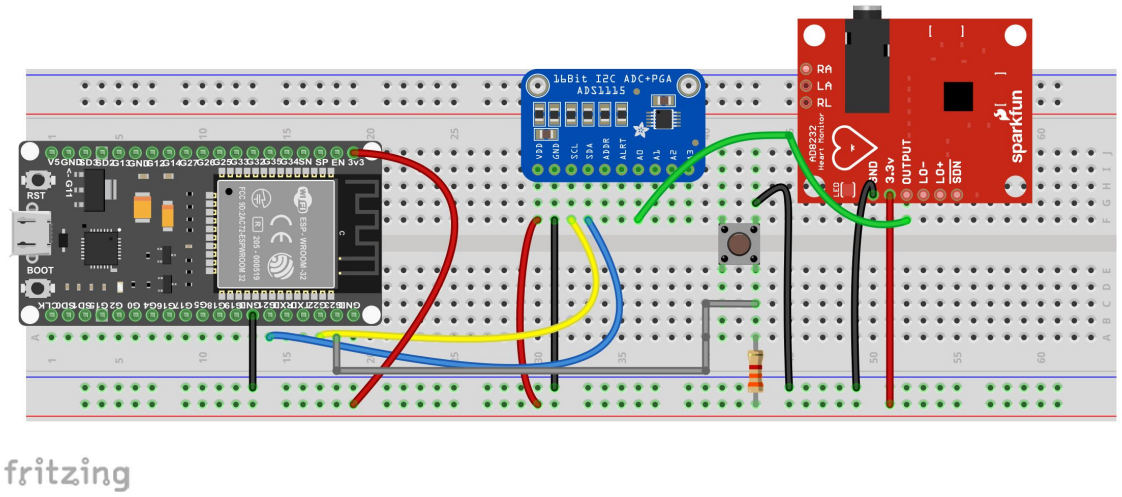
https://www.hackster.io/Yukio/coffee-disease-classification-with-ml-b0a3fc

Forest Fire Detection

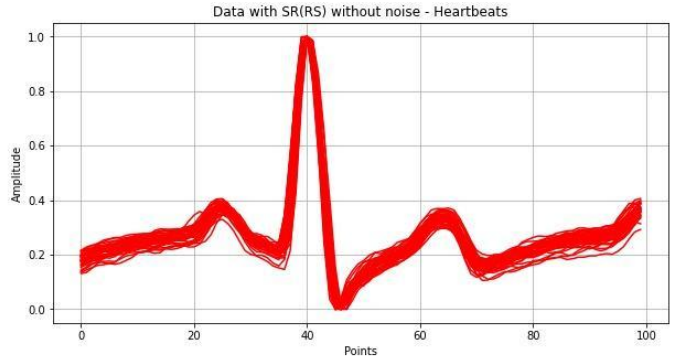
TinyML Aerial Forest Fire Detection

IESTI01 - Forest Fire Detection – Proof of Concept

Health - Human Sensing



fritzing



Guilherme Silva
Engenheiro - UNIFEI

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

Classifying mosquito wingbeat sound using TinyML

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

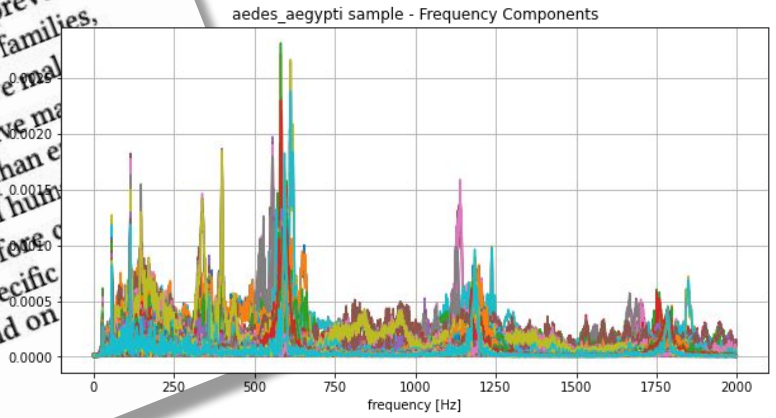
Marcelo Rovai
Universidade Federal de Itajubá
Itajubá, Brazil
rovai@unifei.edu.br

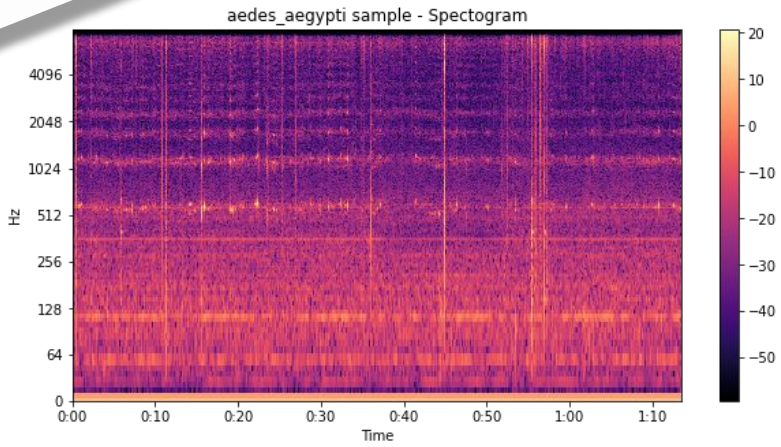
Marco Zennaro
ICTP
Trieste, Italy
mzennaro@ictp.it

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.

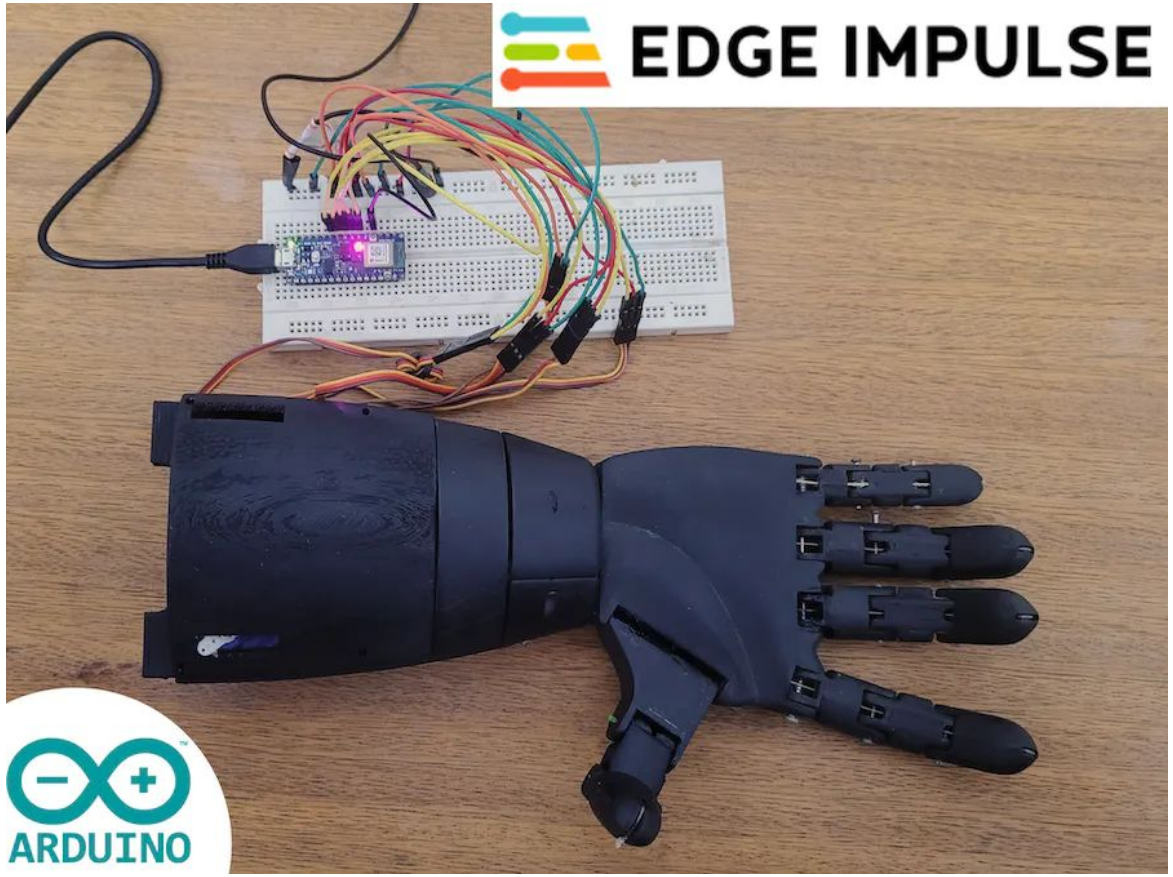
affected. People from poor communities with little access to health care and clean water sources are also at risk. Although anti-malarial drugs exist, there's currently no malaria vaccine. Vector-borne diseases also exacerbate poverty. Illness prevent people from working and supporting themselves and 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

Projects by Students (UNIFEI – IESTI01)

- **Sound:**
  - Earthquake detection
  - Covid Detection (cough)
  - Key Detection
  - Pulmonary Disease
  - Snore Detection
  - Bionic Hand Control

- **Other Sensors:**
  - Bionic Hand – Finger Detection
  - Electric Charges
  - ECG – Fibrial Atrilation detection

- **Image:**
  - Mask Detection
  - Forest Fire Detection
  - Helmet Detection
  - Water Consumption (hydrometer)
  - Sign Language
  - Coffee Disease Classification
  - Bee Counting

- **Vibration:**
  - Personal Trainer
  - Bearing – Anomaly Detection

UNIFEI IESTI01 2021.1 – Final Projects UNIFEI IESTI01 2021.2 – Final Projects UNIFEI IESTI01 2022.1 – Final Projects

TinyML4D **Show & Tell** Presentations

Date	Thread	Video
March 30th, 2023	Thread here when ready	Video here when ready
Feb 23rd, 2023	**thread here**	**https://youtu.be/BAEdiI7X68Y** 3
January 26th, 2023	**thread here** 17	**https://youtu.be/-0xRZ-5UYUc** 9
December 1st, 2022	**thread here** 2	**https://youtu.be/e49pkjnIMIQ** 8
October 27th, 2022	**thread here** 2	**https://youtu.be/s8_hKpOWUwY** 1

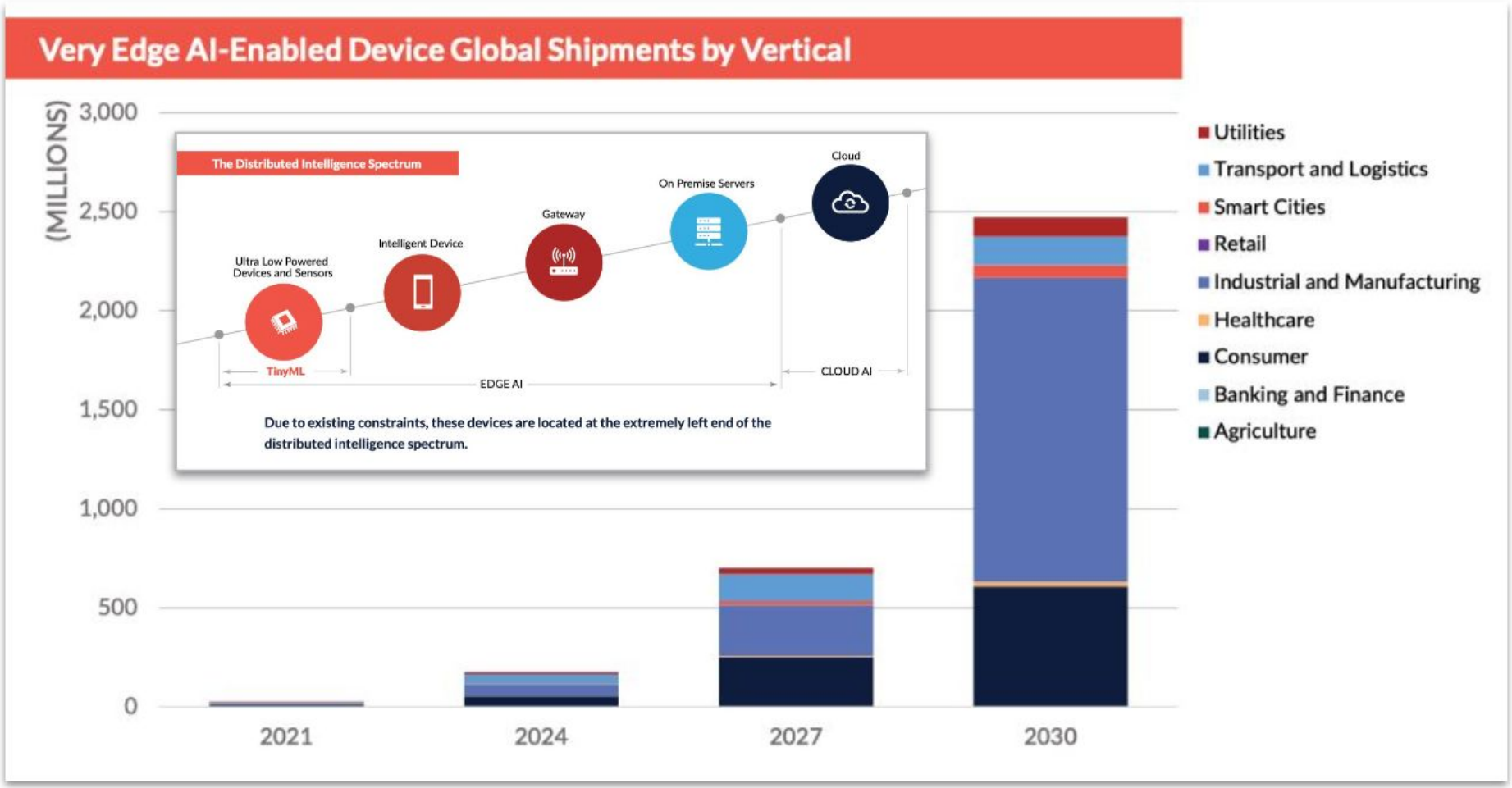
TinymML4D Academic Network Show and Tell Main Index.

The TinyML4D Academic Network Students should use this form to sign up for the latest presentations.
https://forms.gle/ic52HZMqVv4pBrkP7 2

Another site to talk about the presentations is on the Tiny Machine Learning Discord channel at https://discord.gg/rPV6m3wE 4 or at https://discord.gg/zKWgwhSAEY

The Future of the EdgeAI

Massive Potential for Impact



Source: ABI Research: TinyML

Edge AI has a very high impact potential, and it is for now!

Conclusion

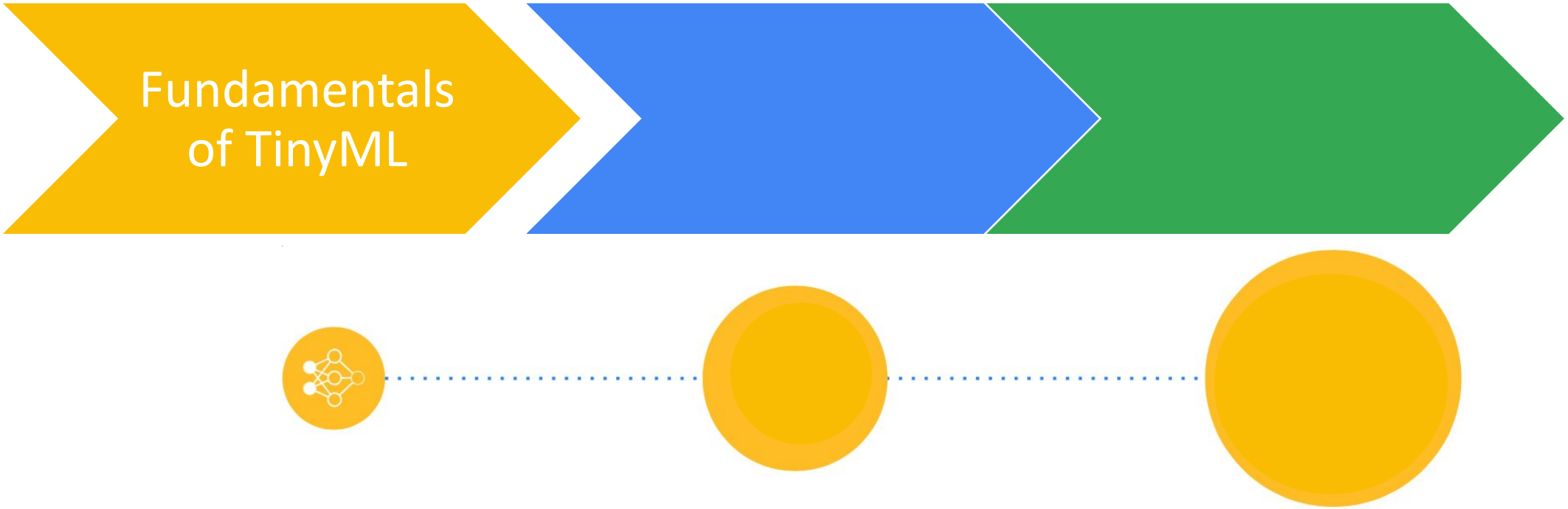
The Future of ML is Tiny and Bright

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

What will We learn?

What will We learn?

Part 1

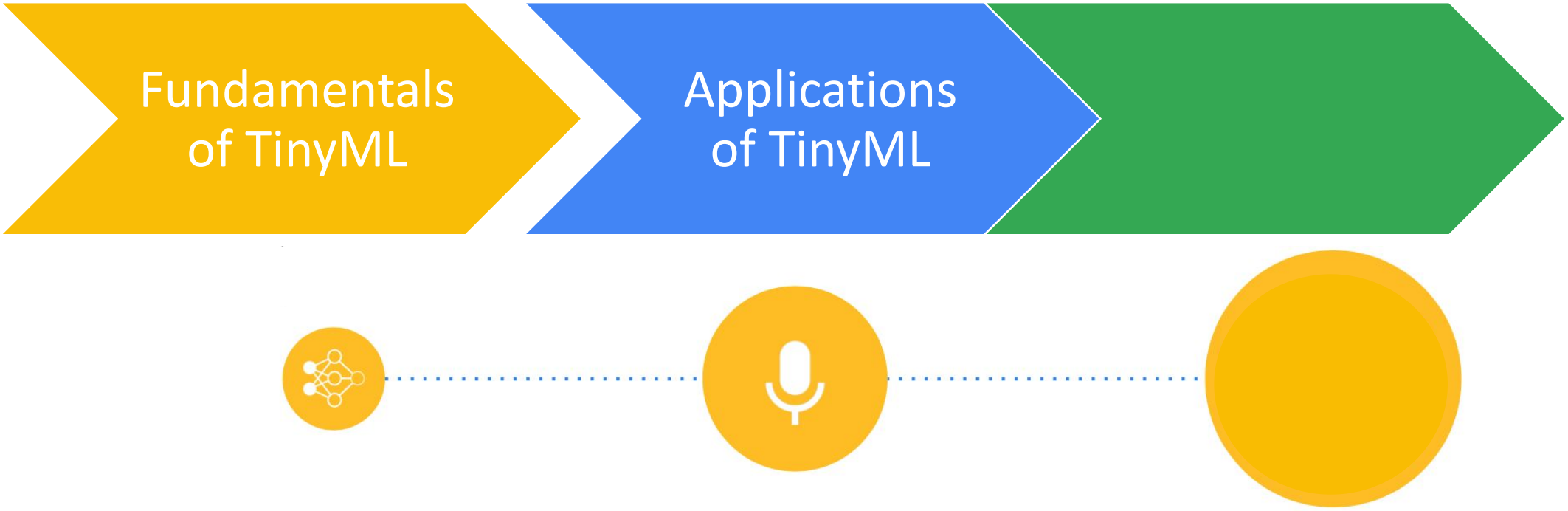


Part 1 is all about talking about what is the language of machine learning

What we will learn?

Part 1

Part 2



In Part 2, we will get a sneak peek into the variety of different TinyML applications, as keyword spotting (“Alexa”), gesture recognition, understand how to leverage the sensors, and so forth.

What we will learn?

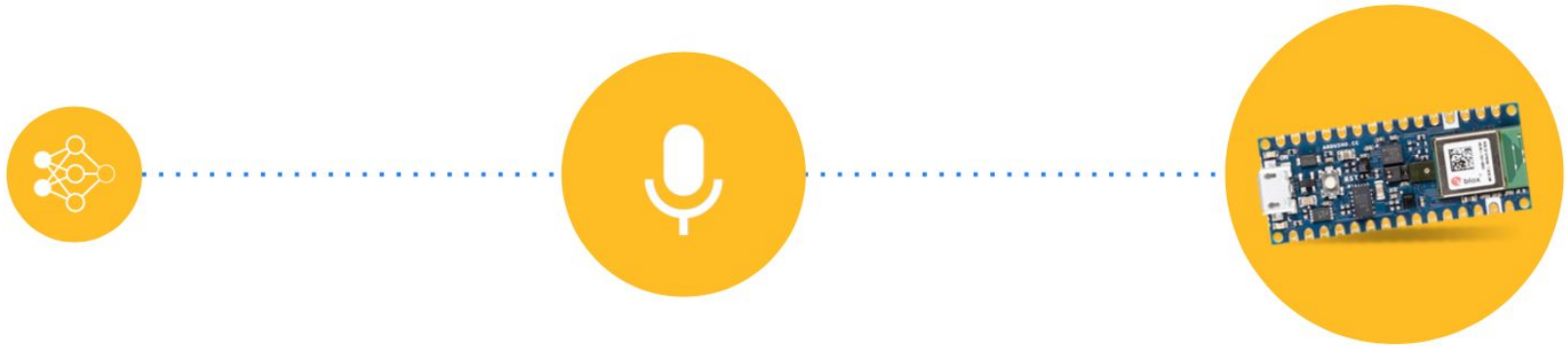
Part 1

Part 2

Fundamentals of TinyML

Applications of TinyML

Deploying TinyML



In Part 2, we will **also** learn how to deploy models on a real microcontroller. Along the way we will explore the challenges unique to and amplified by TinyML (e.g., preprocessing, post-processing, dealing with resource constraints).

Background Requirements

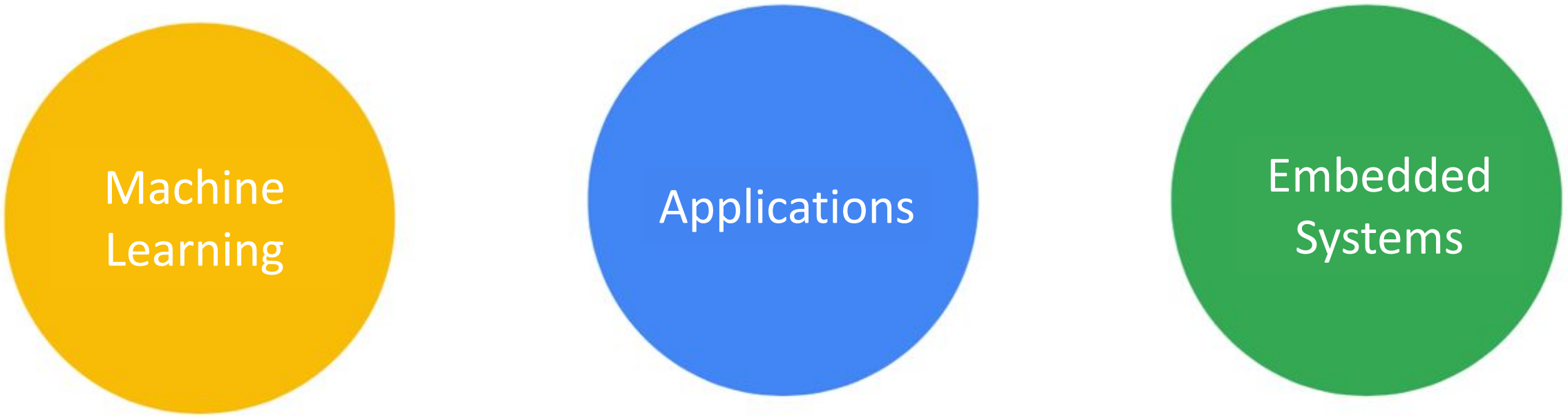
- Python
- TensorFlow
- Google Colab
  - Jupyter Notebook

- Python
- TensorFlow (Lite)
- Google Colab
- Edge Impulse Studio

- Python
- TensorFlow (Lite-Micro)
- Google Colab
- Edge Impulse Studio
- IDE (as Arduino)
- C/C++

This course combines **computer science** with **engineering** to feature real-world application case studies that examine the challenges facing **TinyML deployments**.

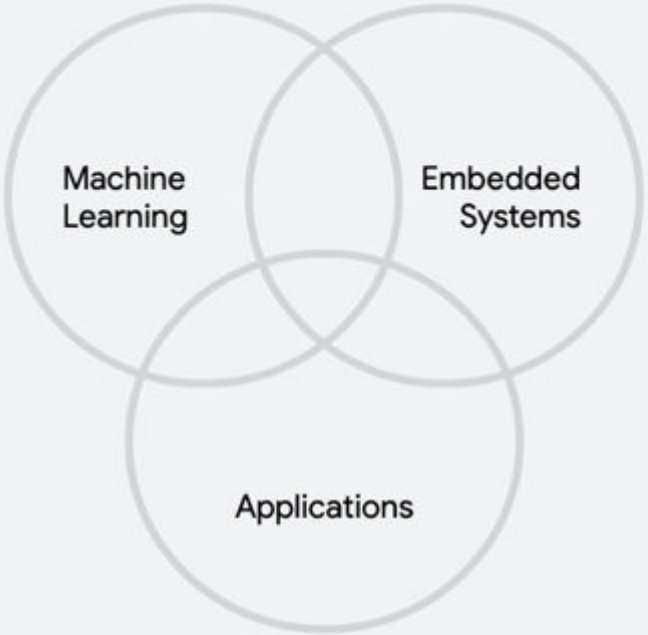
What areas we will learn?



We will learn the **fundamentals of each of these areas**, just enough to focus on the goal of being able to build **TinyML applications**.

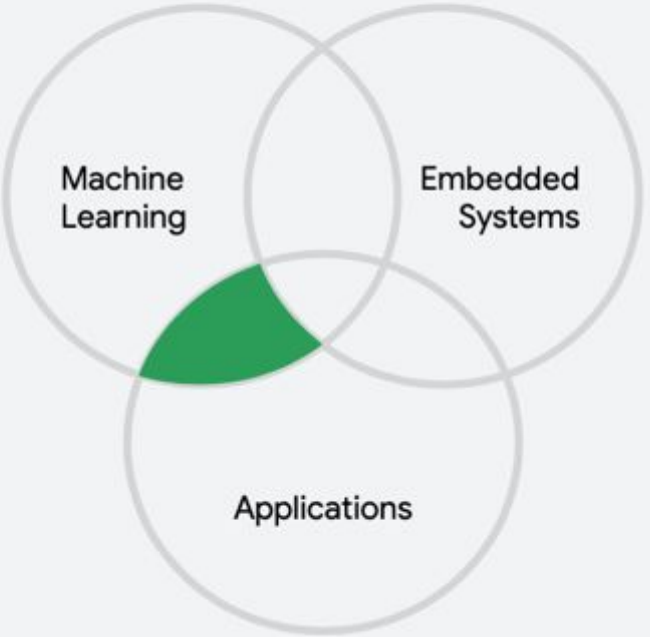
Interactions

In addition, we will bring these diverse topics together to reveal the interesting learnings at the various **intersections**



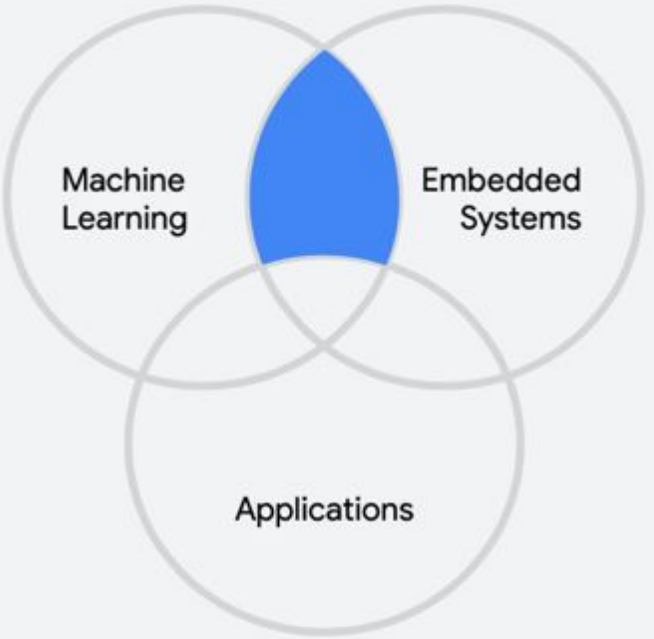
Interactions

How **machine learning** can enable new and interesting **TinyML applications**?



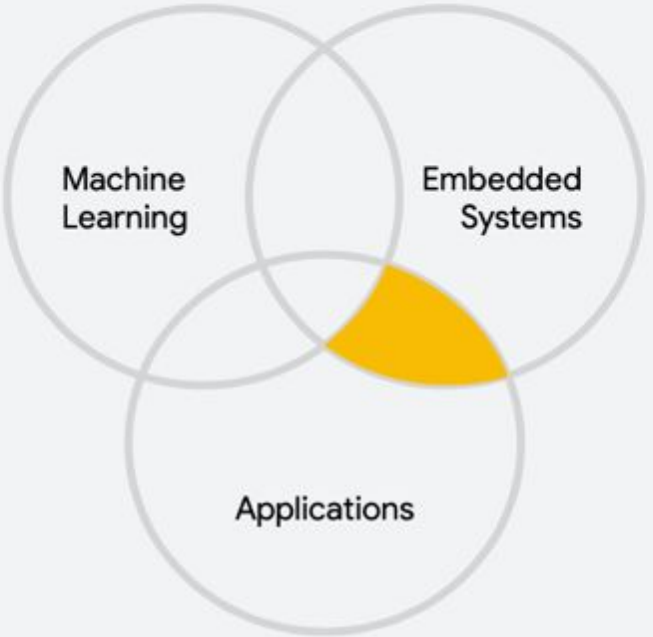
Interactions

What are the **challenges** with enabling **machine learning** on **tiny**, resource-constrained **embedded devices**?



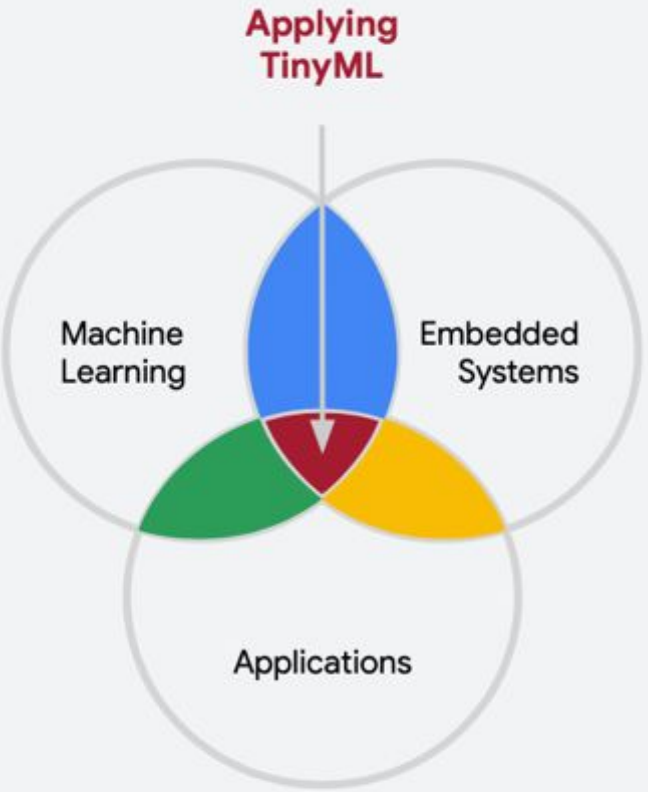
Interactions

What type of new **use cases** can we possibly enable on **embedded systems** that we could not otherwise do before?



At the End of the Day

Given your understanding of things at these various intersections, you will have a deep understanding for **how to apply TinyML**



How are we going to get there?

Hands-on Learning

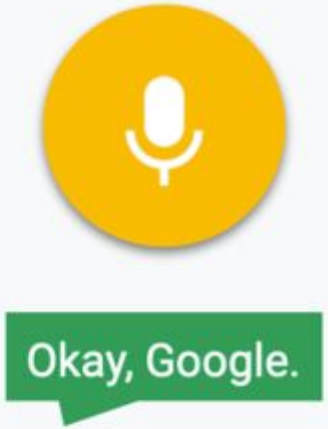
- Software
  - Machine Learning (TensorFlow)
  - Programming environments (Google Colab or Jupyter)
  - Edge Impulse Studio
- Hardware
  - Arduino Nano 33 BLE Sense
  - Sensors

TinyML Arduino Kit

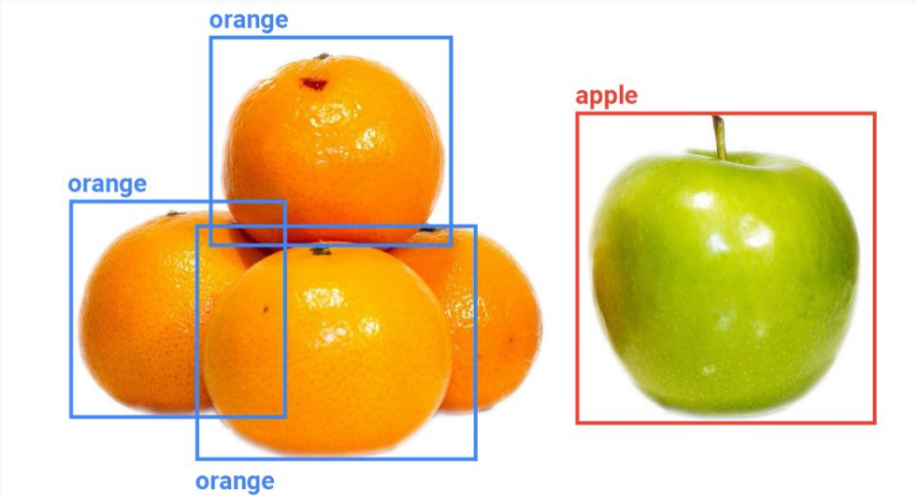
Wio Terminal Kit

Hands-on Activities

Speech



Vision

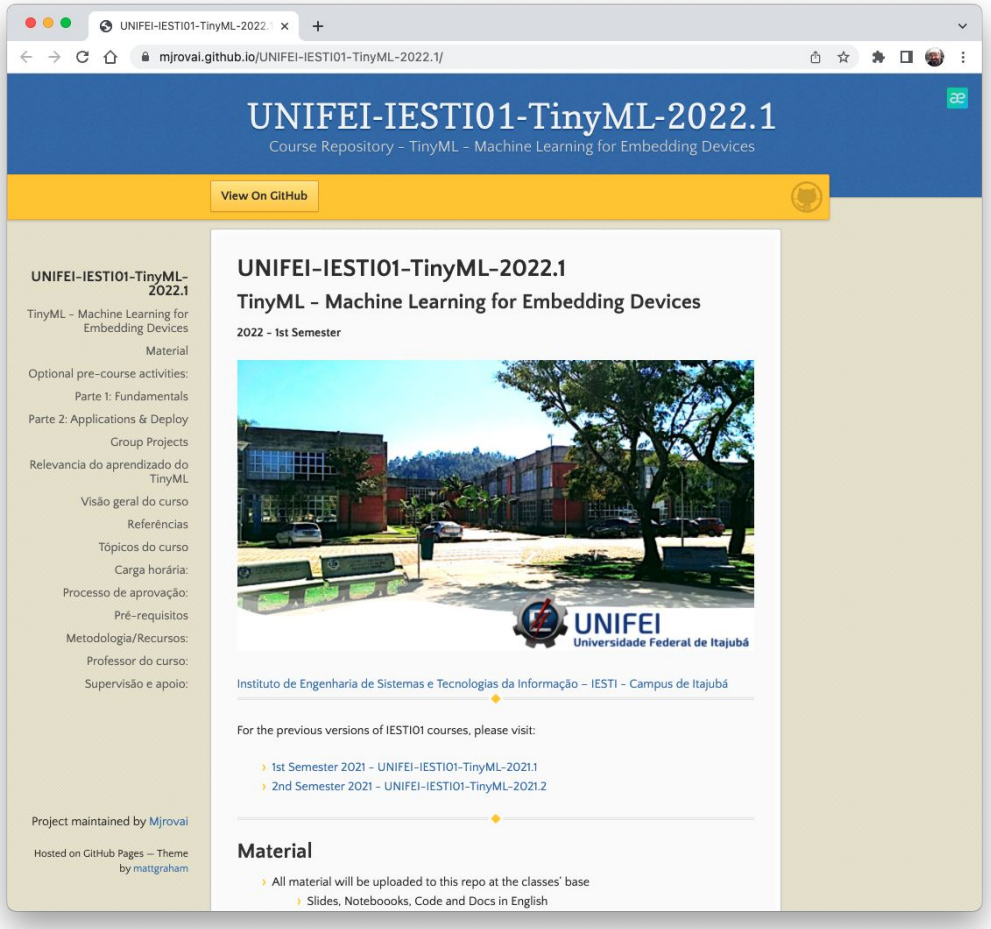


IMU

How is the course structured?

Course Structure

- Weekly video-recorded lectures (15 weeks)
  - Slides
  - Hands-on coding (by teacher & students)
- Weekly Additional Readings
- Possible Guest Lecturer (previous *)
- Assignments
  - Quizzes (Weekly)
  - Notebooks with codes (4)
  - Hands-on lab reports (4)
- Final Project (Groups of 3 or 4 students)
  - Report
  - Presentation



(*) IESTI01 2021.1 -> Daniel Situnayaki, Edge Impulse - US
IESTI01 2021.2 -> Dr. Marco Zennaro, ICTP - Italy

Previous IESTI01 Courses available for consultation

Class planning and approval process

- Minimal suggested Workload (4 hours per week):
  - 30 hours (Weekly recorded classes of about 2h, for 15 weeks)
  - 15 hours of assignments/coding/labs
  - 15 hours in research, individual studies, and final project (in a group)

- Approval process:
  - 1st Evaluation:
    - Individual **Quizzes**: 10%
    - Individual **Exercise Lists** (Notebooks): 25%
    - Group **Project Proposal**: 15%
  - 2nd Evaluation
    - Individual **Quizzes**: 10%
    - Individual **Practical Projects** (Lab reports): 25%
    - Group **Project Presentation (*)** and Final Report: 15%

(*) Examples: Harvard CS249r – TinyML - Final Projects

UNIFEI IESTI01 2021.1 – Final Projects

UNIFEI IESTI01 2021.2 – Final Projects

UNIFEI IESTI01 2022.1 – Final Projects

Main references

- Harvard School of Engineering and Applied Sciences - CS249r: Tiny Machine Learning
- 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
- Fundamentals textbook: “Deep Learning with Python” by François Chollet
- Applications & Deploy textbook: “TinyML” by Pete Warden, Daniel Situnayake
- Deploy textbook “TinyML Cookbook” by Gian Marco Iodice
- Deploy textbook "AI at the Edge" book by Daniel Situnayake, Jenny Plunkett

I want to thank Shawn Hymel and Edge Impulse, Pete Warden and Laurence Moroney from Google, Professor Vijay Janapa Reddi and Brian Plancher from Harvard, and the rest of the TinyMLedu team for preparing the excellent material on TinyML that is the basis of this course at UNIFEI.

The IESTI01 course is part of the **TinyML4D**, an initiative to make TinyML education available to everyone globally.

Thanks

UNIFEI