Embedded ML (TinyML) Intro & Applications

Marco Zennaro, PhD

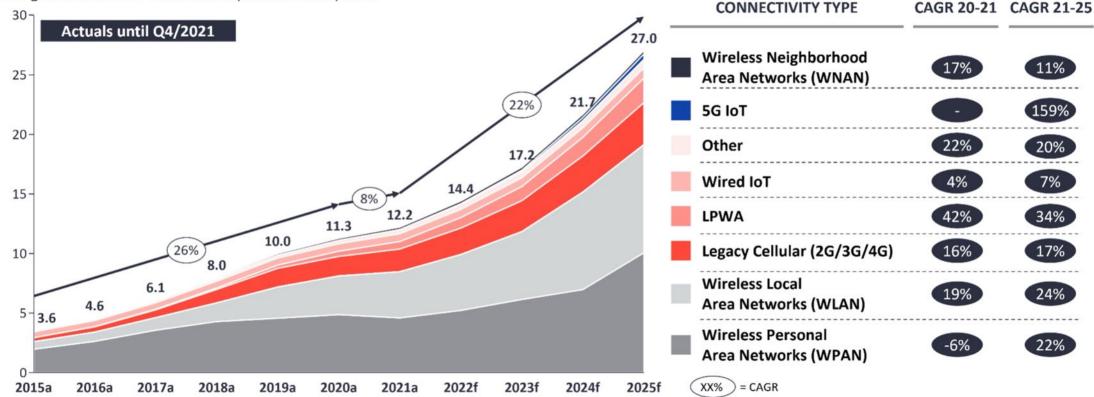
TinyML4D Academic Network Co-Chair



Internet of Things (IoT)

Global IoT Market Forecast [in billion connected IoT devices]

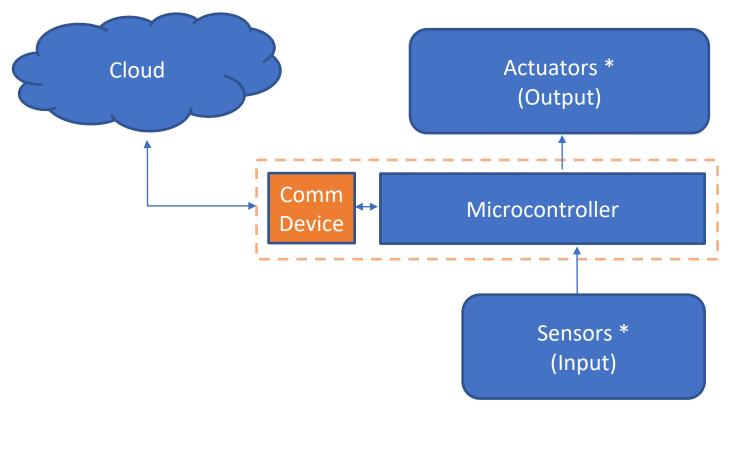
Number of global active IoT Connections (installed base) in Bn



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 onedirectional 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.

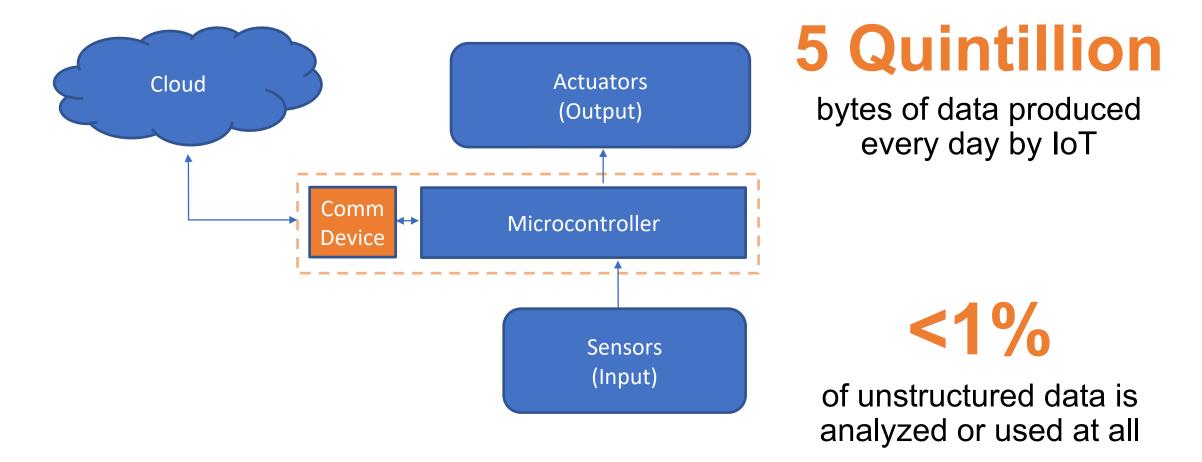
Typical IoT Project



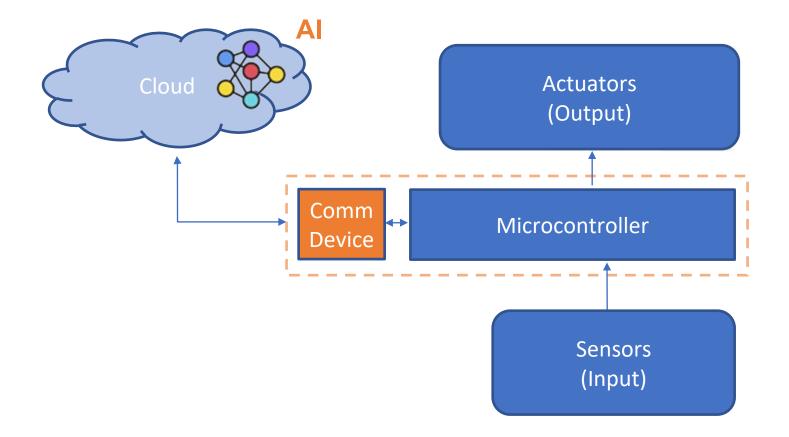




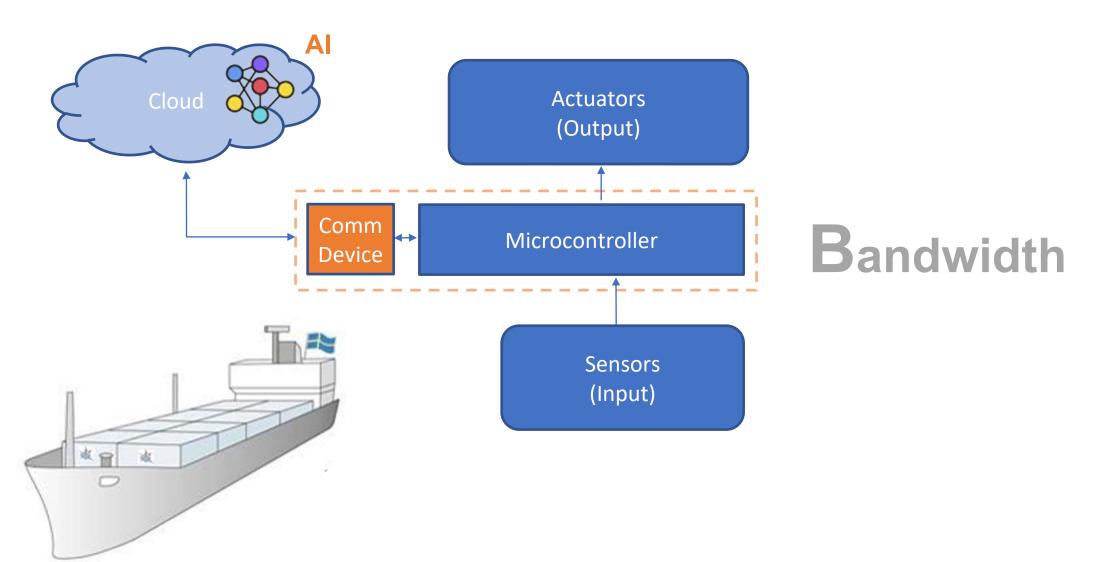
Typical IoT Project



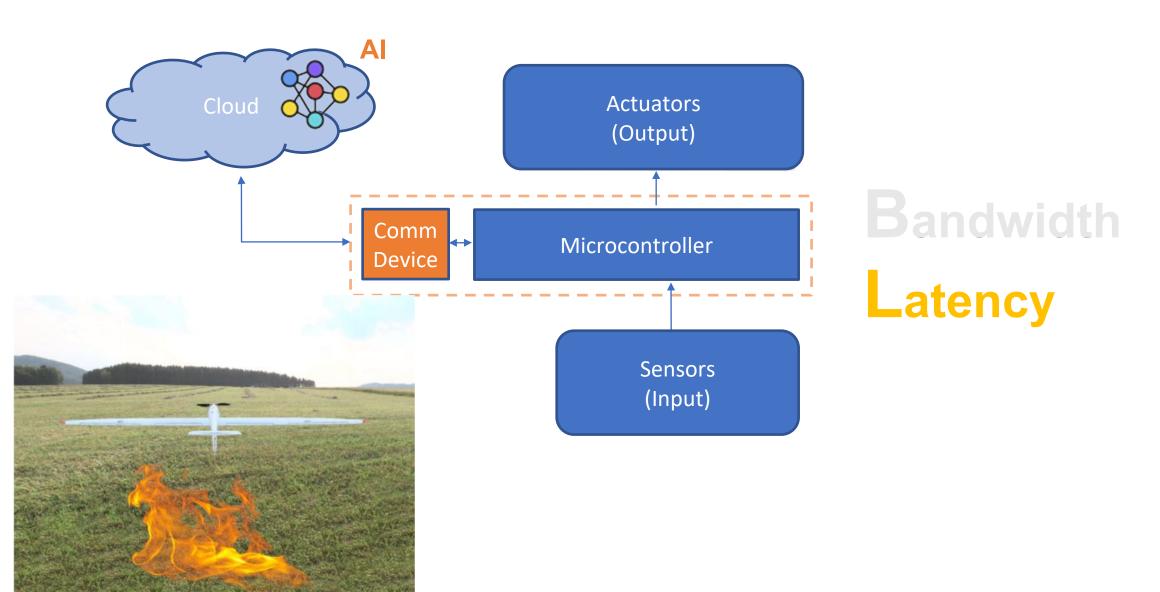
Source: Harvard Business Review, <u>What's Your Data Strategy?</u>, April 18, 2017 Cisco, <u>Internet of Things (IoT) Data Continues to Explode Exponentially. Who Is Using That Data and How?</u>, Feb 5, 2018



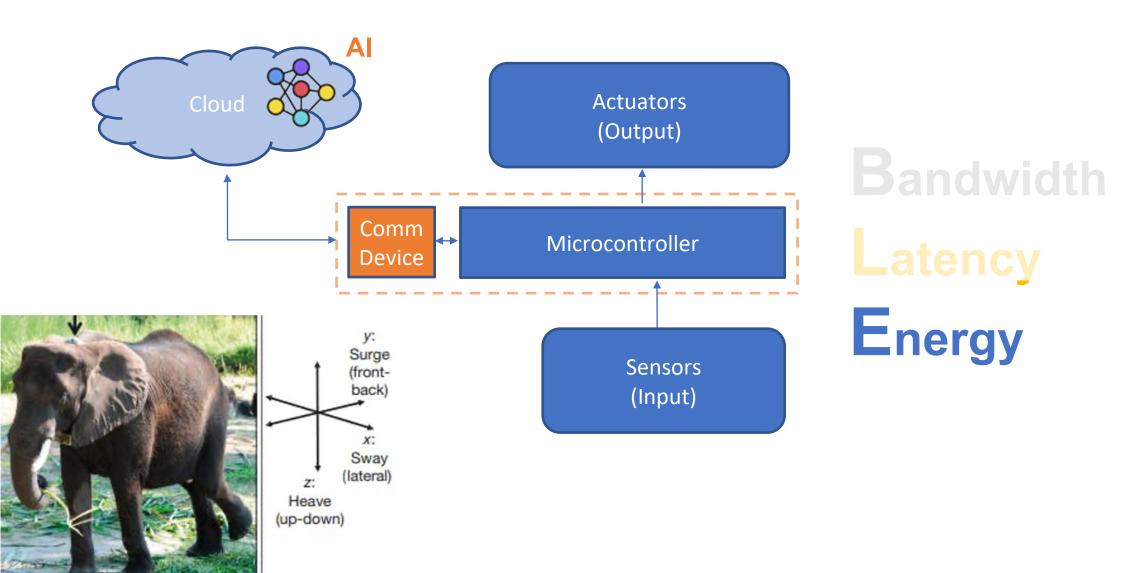




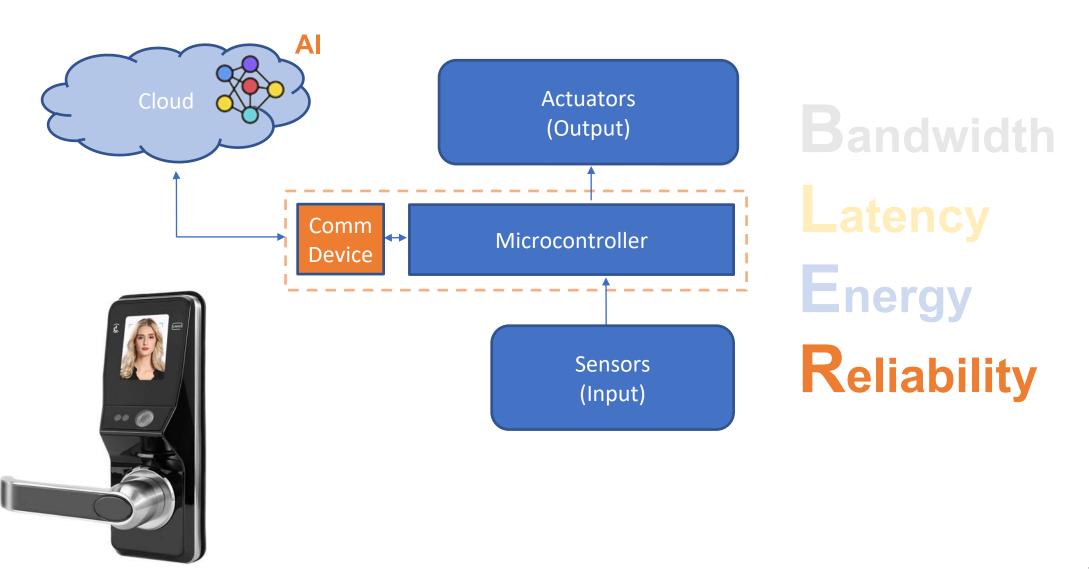




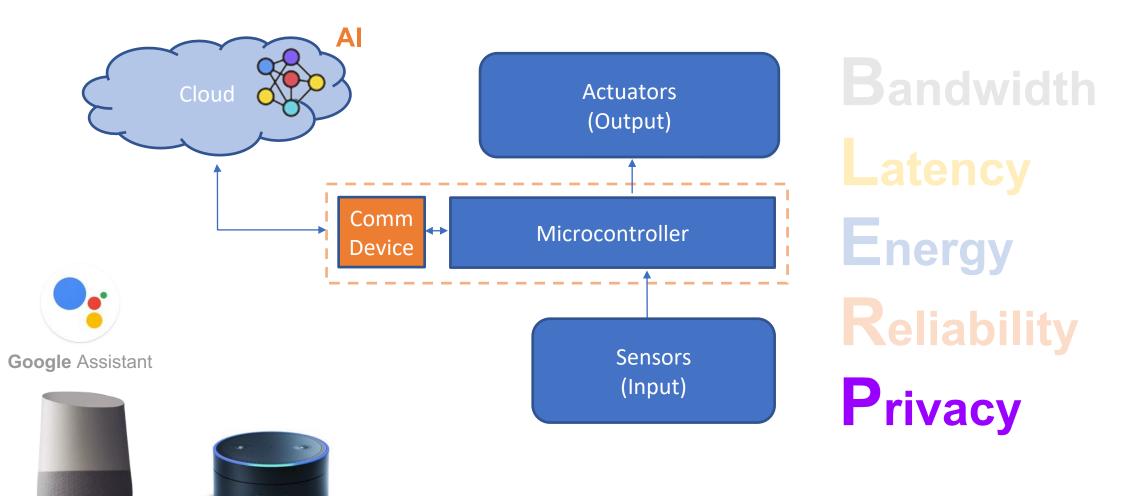




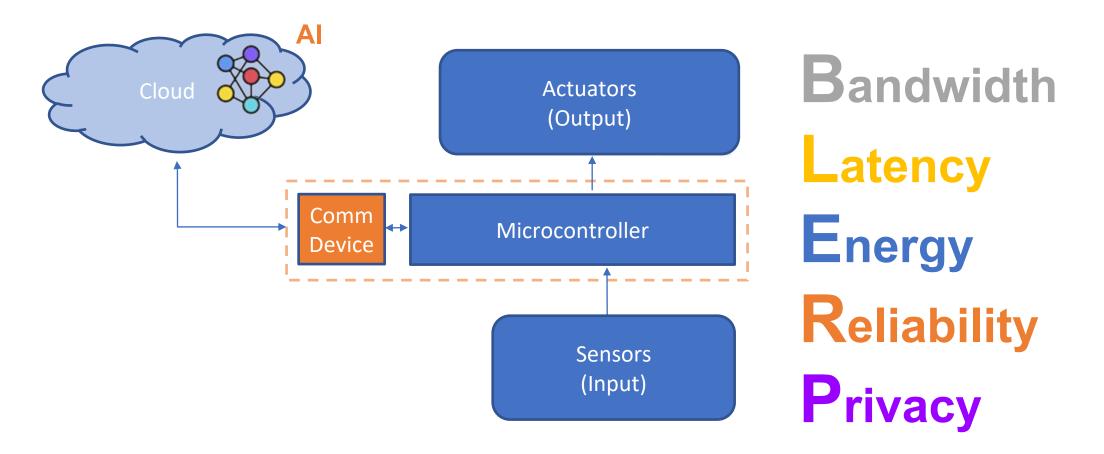






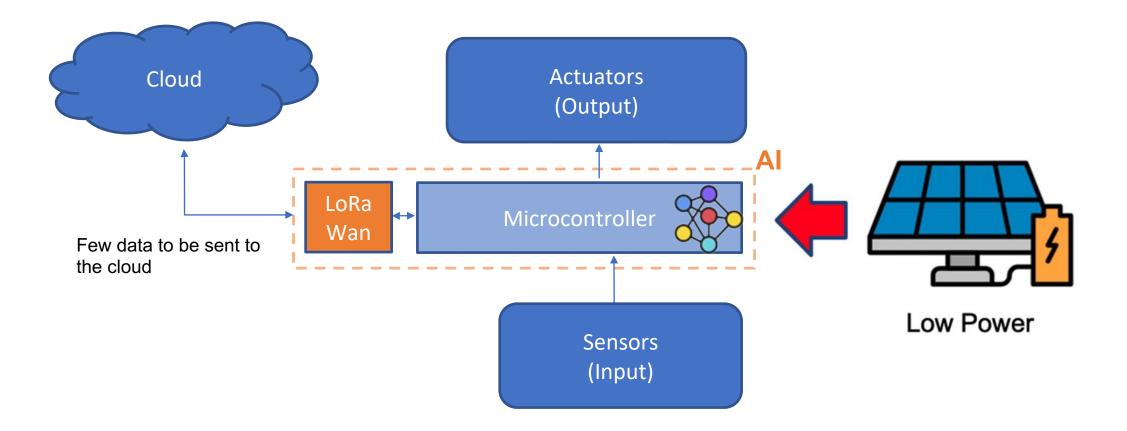






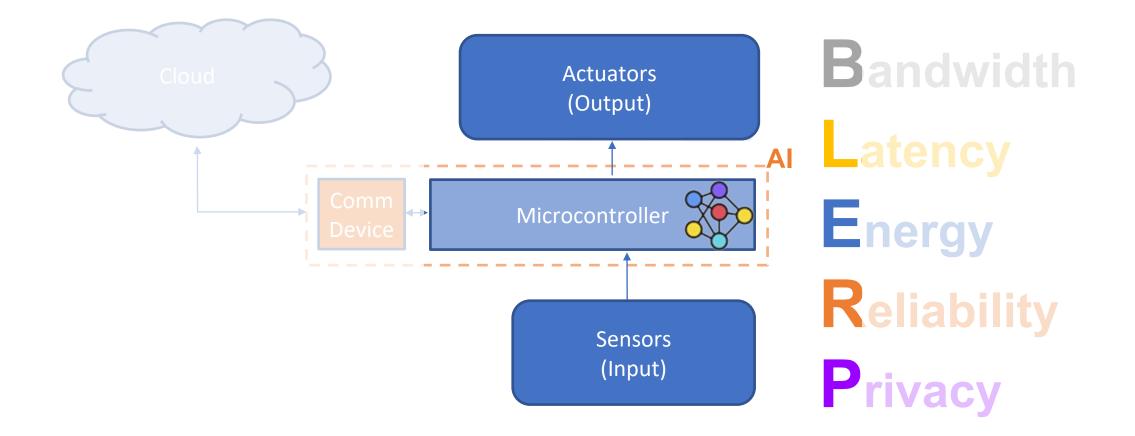
... Solution ?

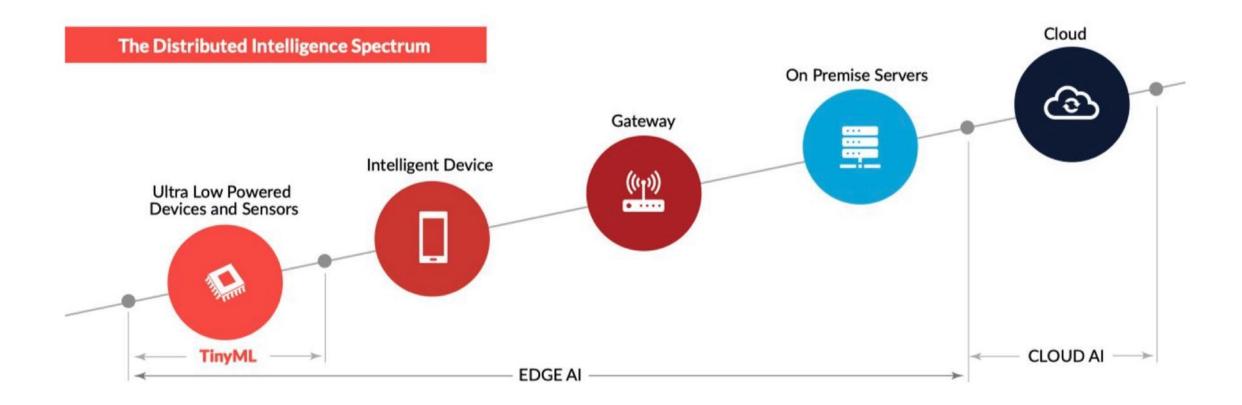
IoT 2.0 * - Edge AI/ML * Intelligence of Things



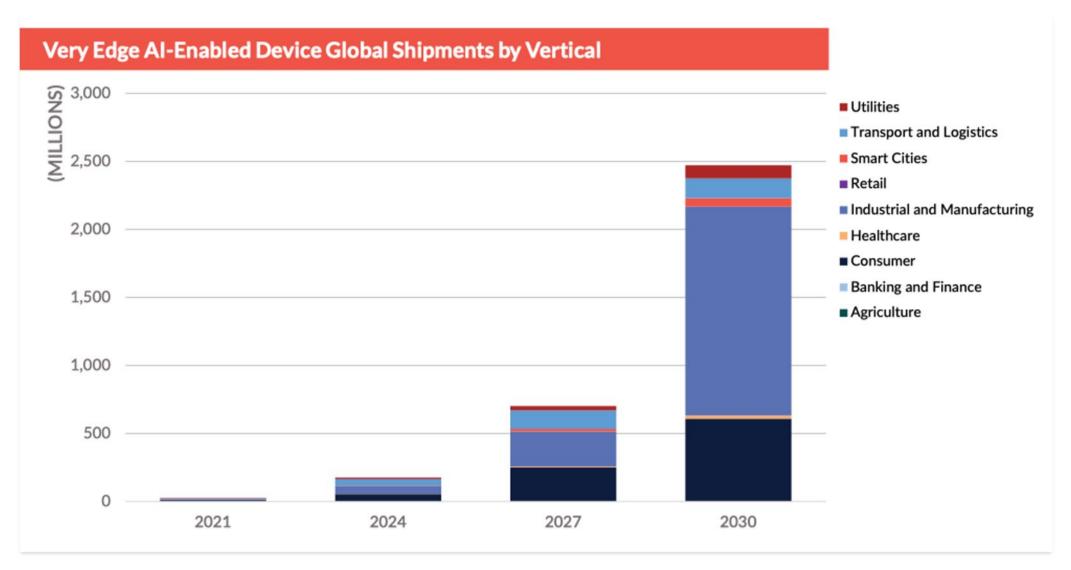
... Solution -> ML goes close to data

When to use an Edge AI/ML approach:

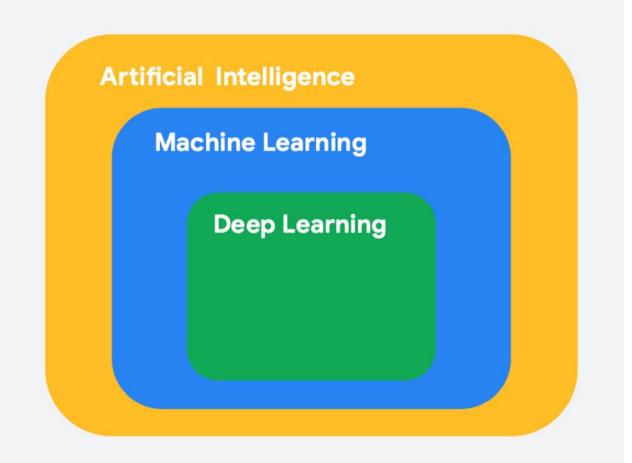




Market Forecast



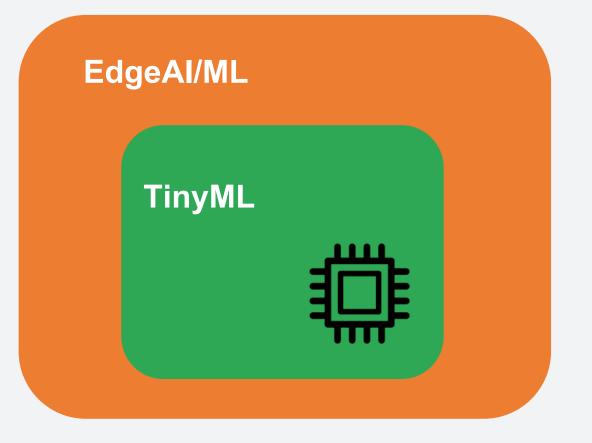
Embedded ML (TinyML) Introduction



Al: Any technique that enables computers to mimic human behavior

ML: Ability to learn without explicitly being programed

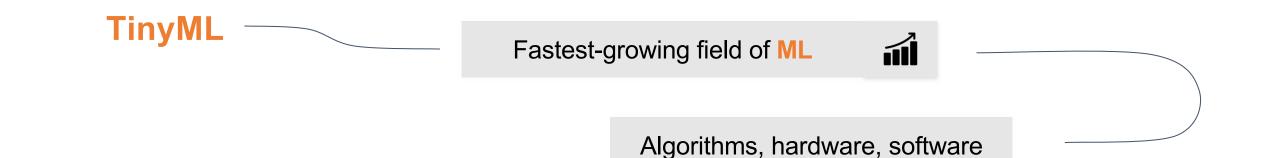
DL: Extract patterns from data using neural networks

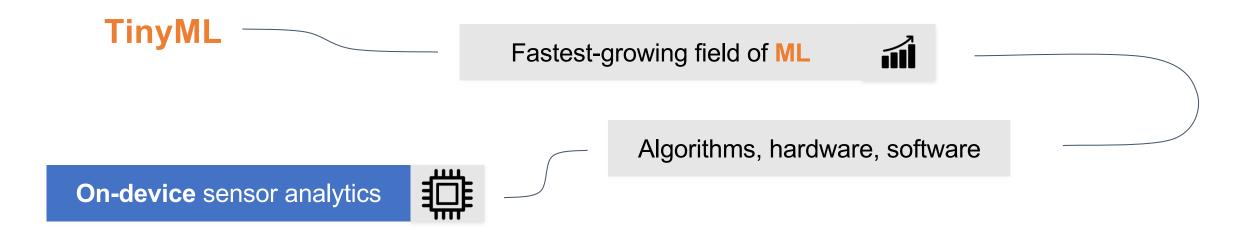


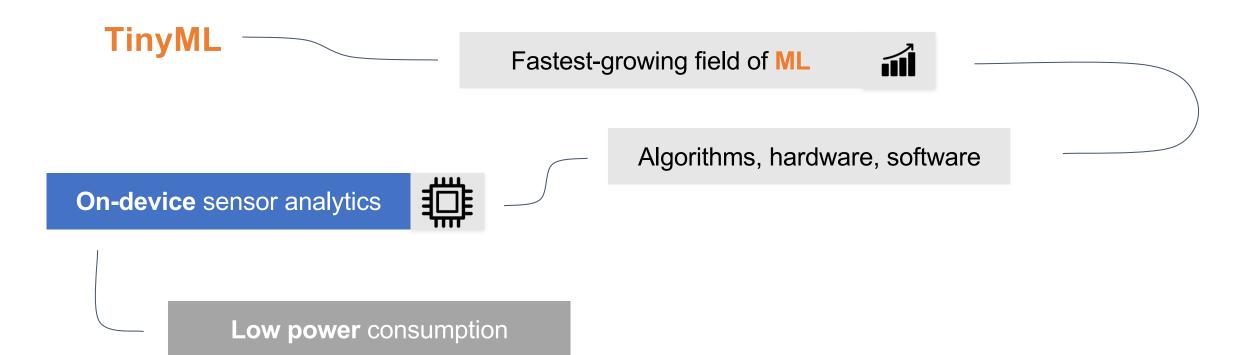
Edge Al (or Edge ML) is the processing of Artificial Intelligence algorithms on edge, that is, on users' devices. The concept derives from **Edge Computing**, which starts from the same premise: data is stored, processed, and managed directly at the Internet of Things (IoT) endpoints.

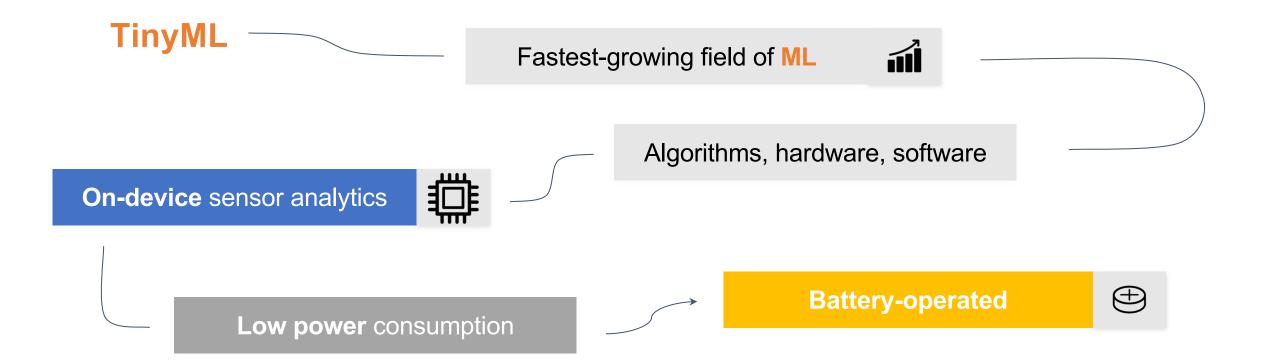
TinyML is a subset of EdgeML, where sensors are generating data with ultra-low power consumption (batteries), so that we can ultimately deploy machine learning continuously ("always on devices")

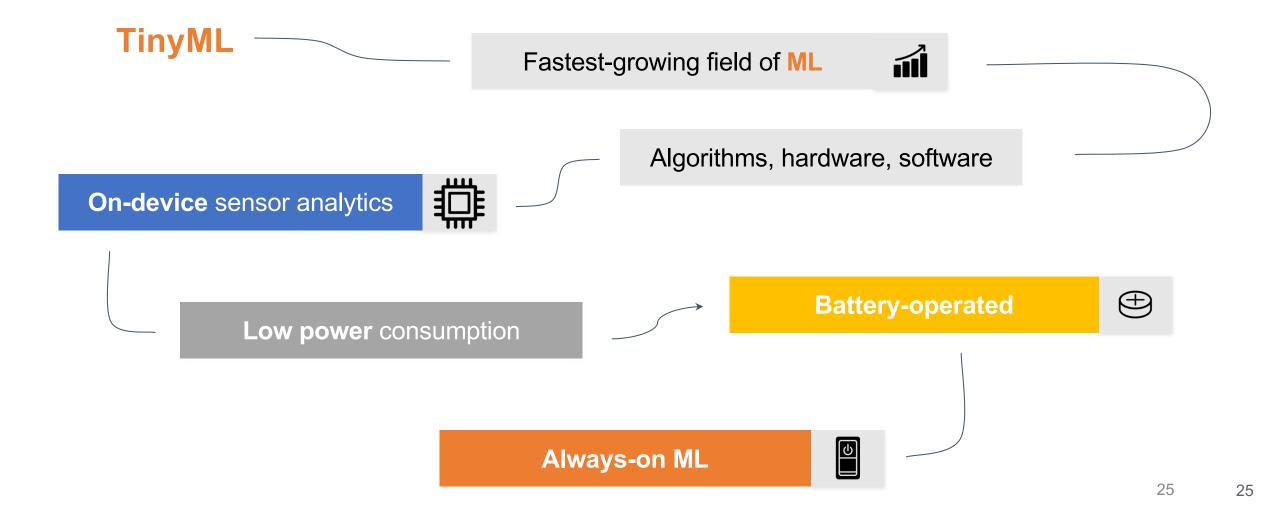




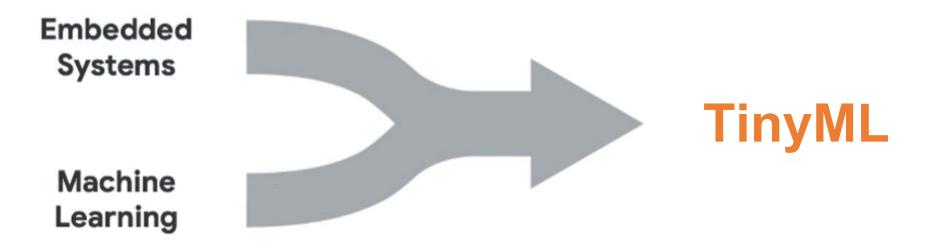




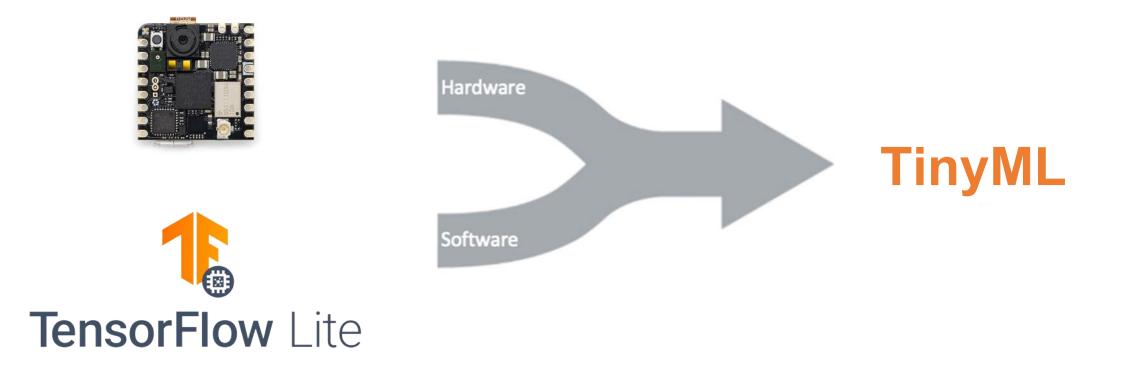




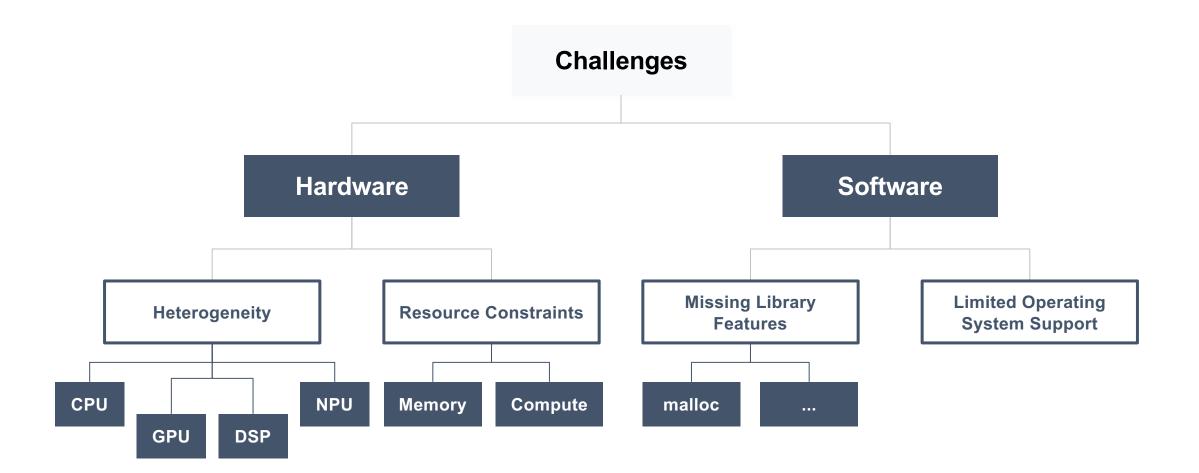
What Makes TinyML ?

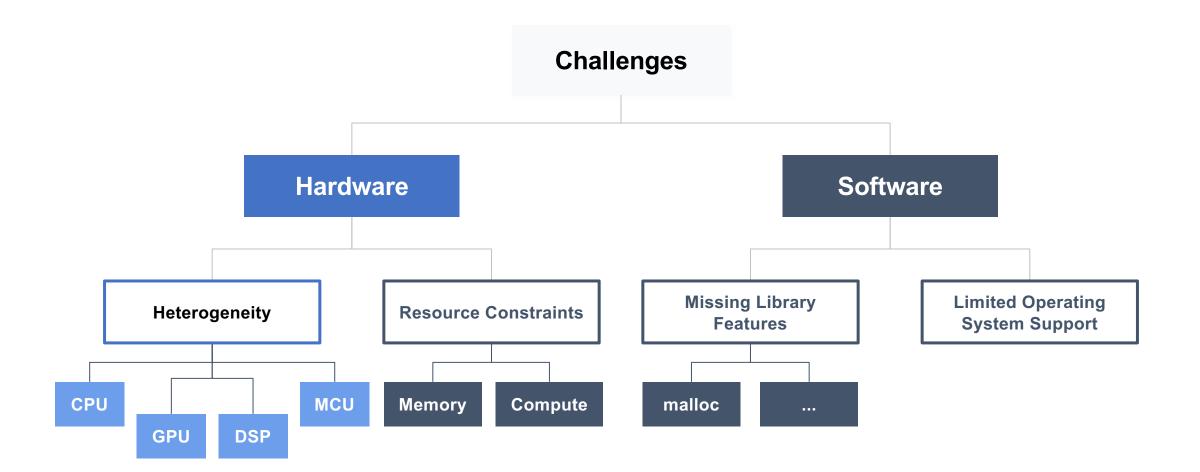


What Makes TinyML ?



TinyML Challenges





250 Billion *MCUs today*

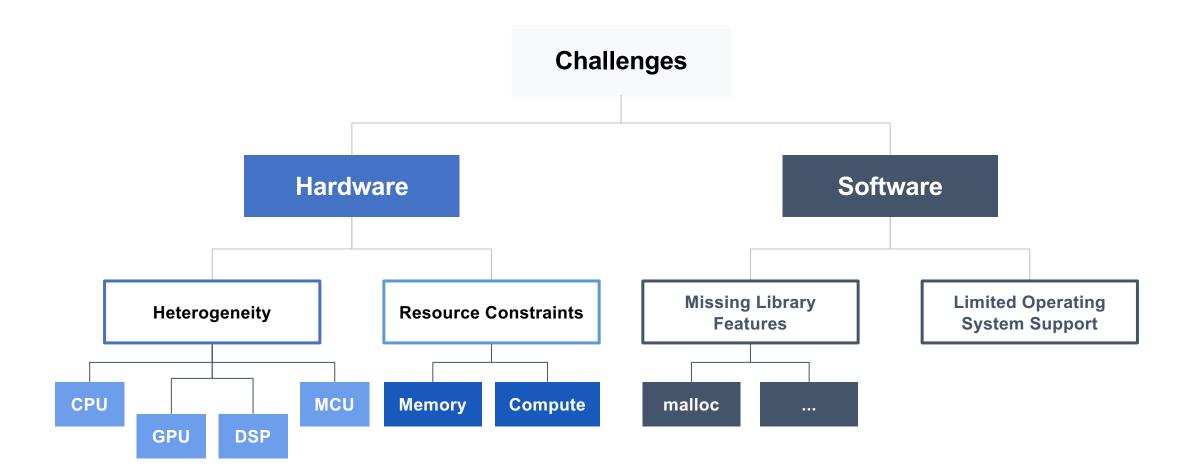
Hardware



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	\$	\$\$\$	\$	\$\$	\$\$\$\$

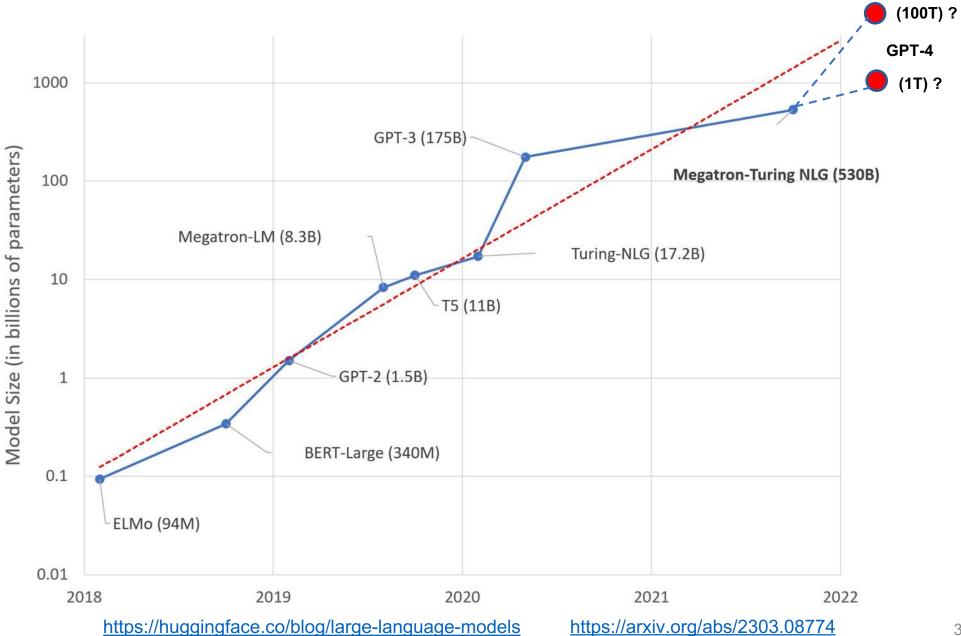


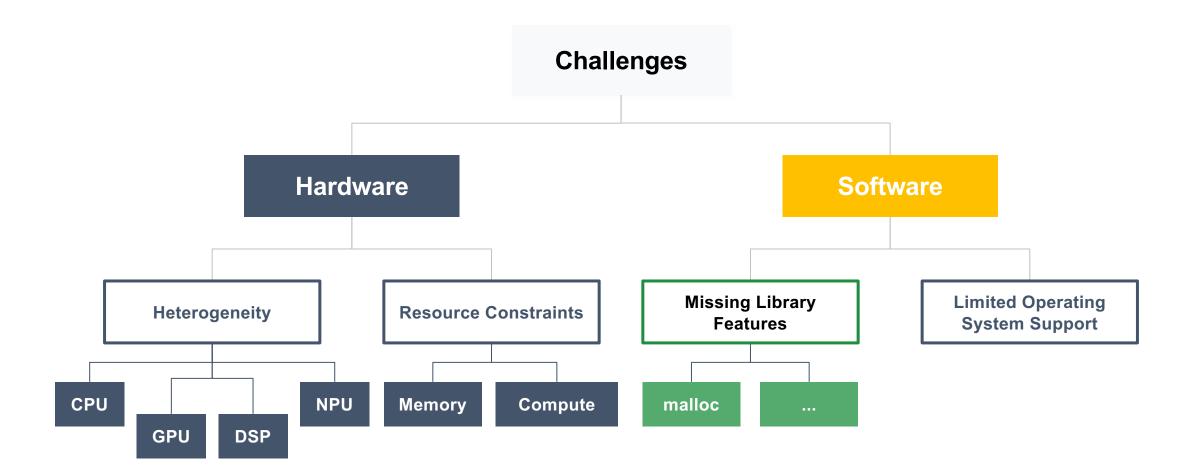
Hardware

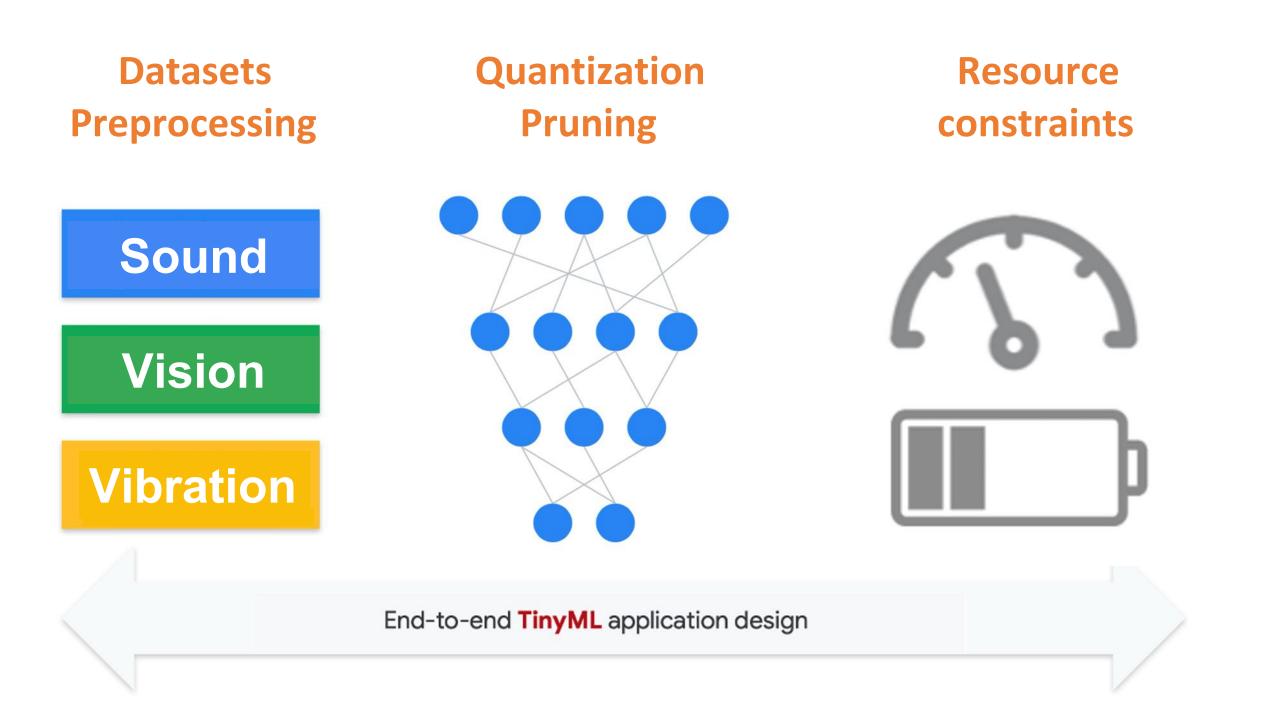


	Raspberry Pico (W)	Arduino Nano Sense	ESP 32	Seeed XIAO Sense / ESP32S3	Arduino Pro
32Bits CPU	Dual-core Arm Cortex-M0+	Arm Cortex-M4F	Xtensa LX6 Dual Core	Arm Cortex-M4F (BLE) Xtensa I X7 Dual Core	Dual Core Arm Cortex M7/M4
CLOCK	133MHz	64MHz	240MHz	64 / 240MHz	480/240MHz
RAM	264KB	256KB	520KB (part 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	\$	\$\$\$	\$	\$\$	\$\$\$\$

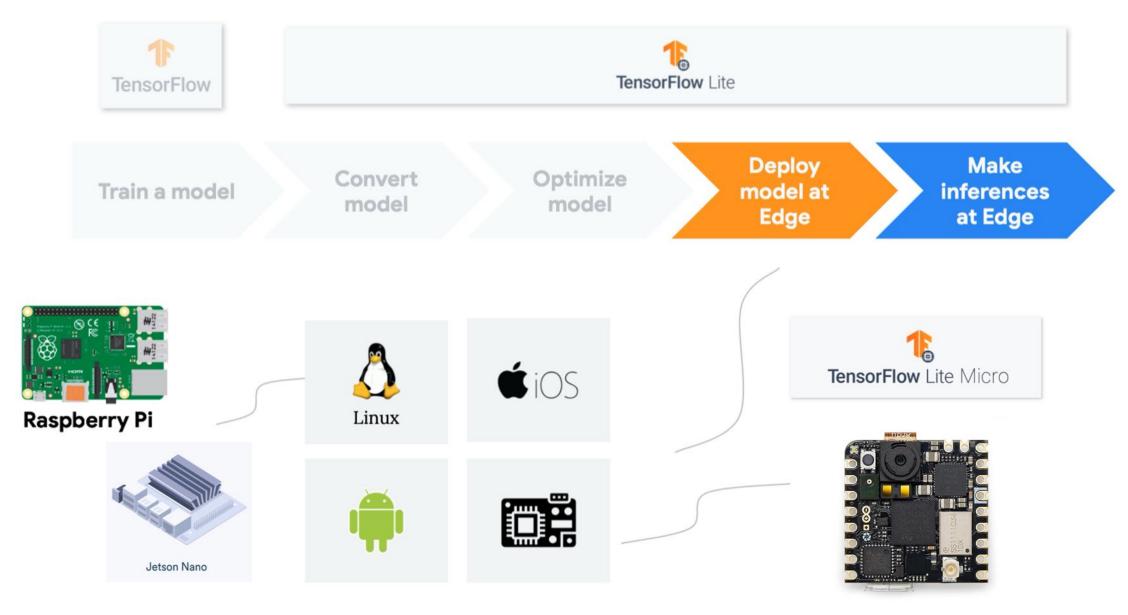
https://media.digikey.com/Resources/Maker/the-original-guide-to-boards-2022.pdf

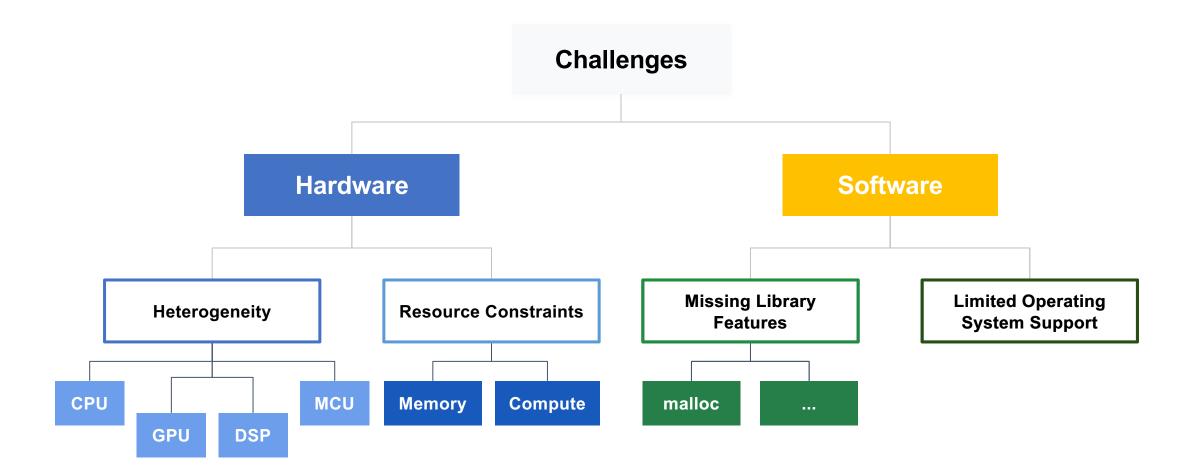






Software



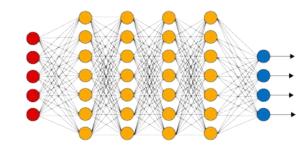


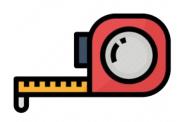
Power EdgeML Application Complexity vs. HW **TinyML** Video Classification 2 MB+ **Object Detection Complex Voice** Processing 1 MB+ Image Classification 250 KB+ **KeyWord Spotting** Audio Classification 50 KB **Anomaly Detection** Sensor Classification XIAO 20 KB ESP32 Arduino Pro Jetson Nano SmartPhone **Rpi-Pico** Arduino Nano RaspberryPi (Cortex-M7) (Cortex-A + GPU) (Cortex-M0+) (Cortex-M4) (Cortex-A) CPU Power / Memory

Application Complexity

How to Train a ML Model?

Machine Learning Workflow ("What")

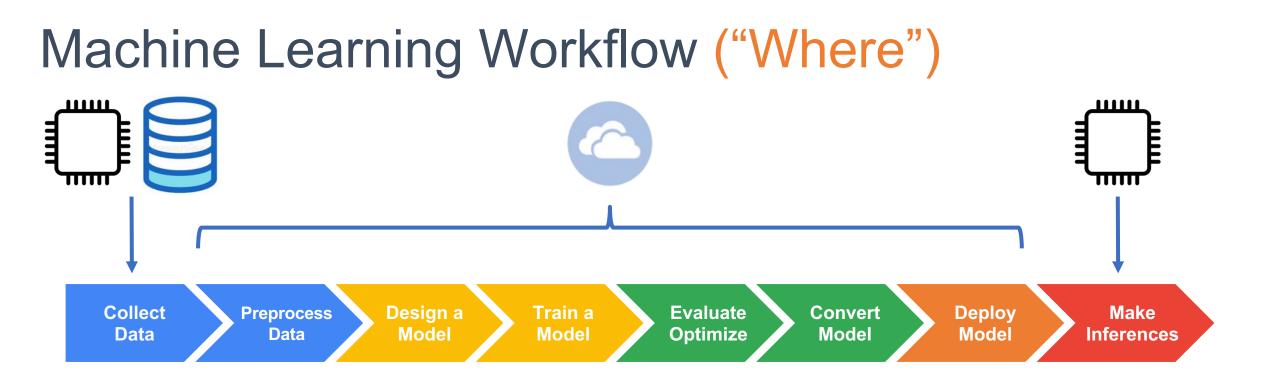


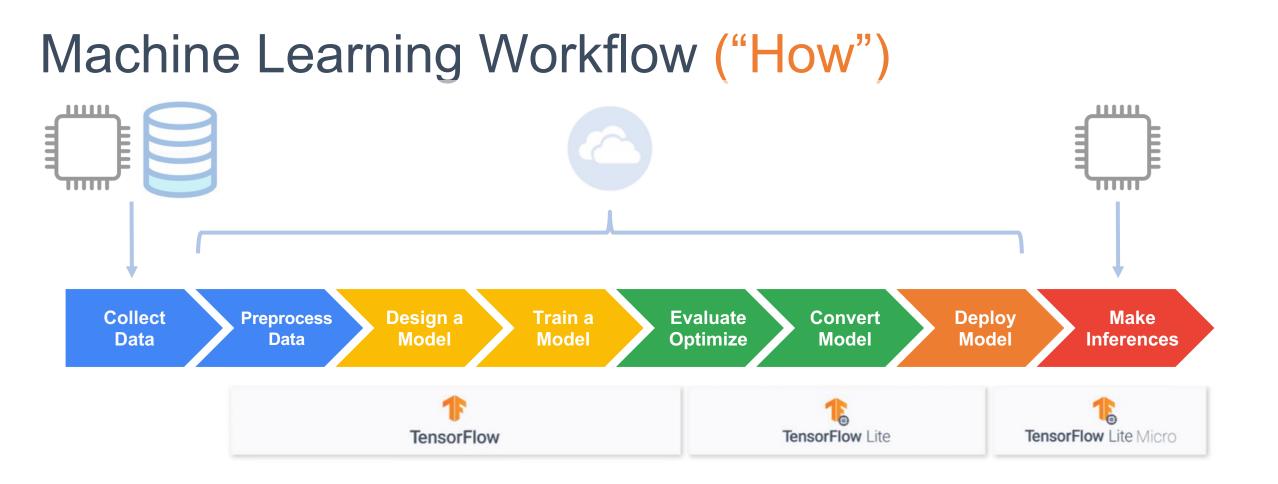


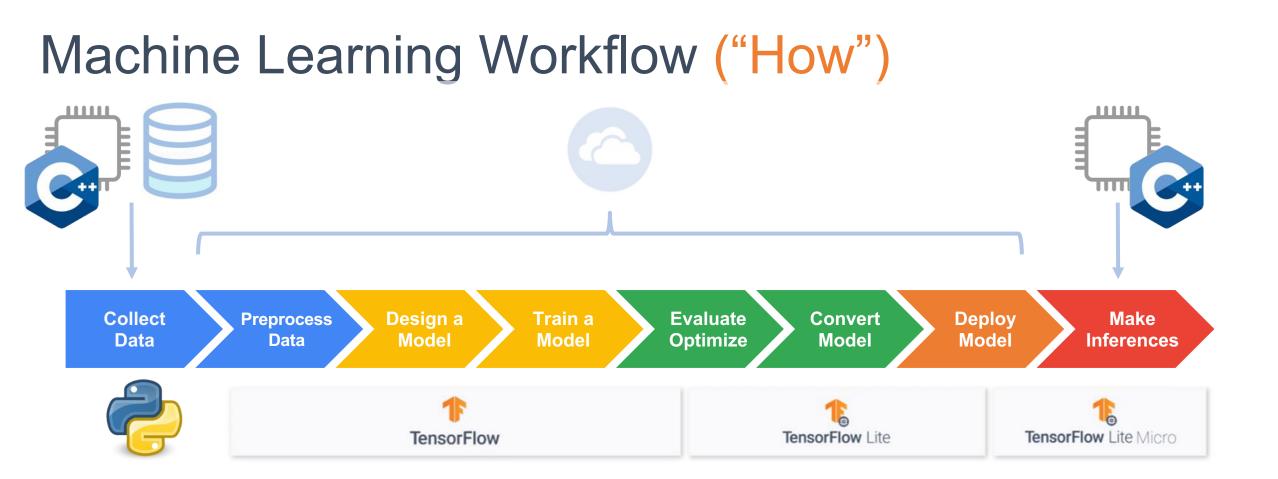


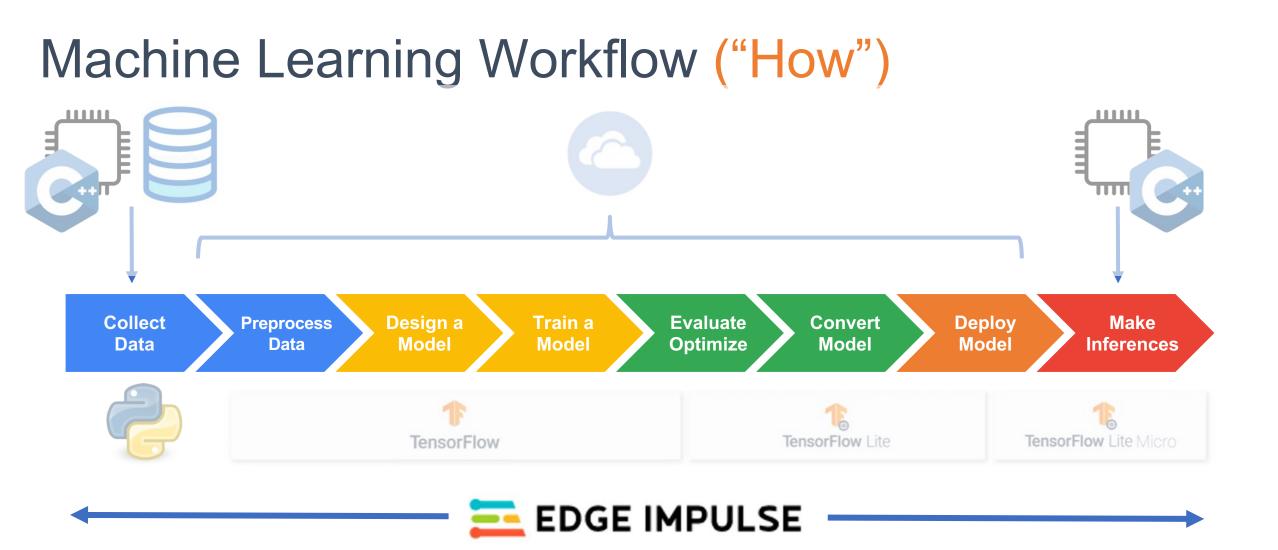




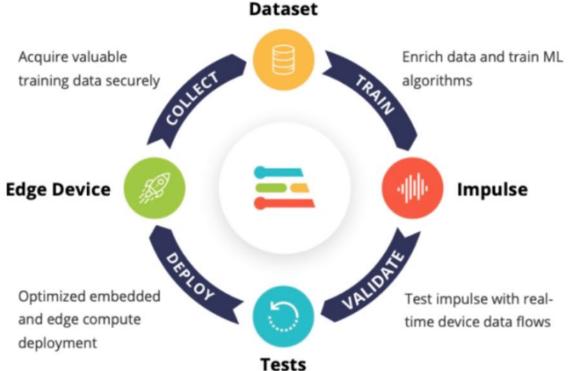


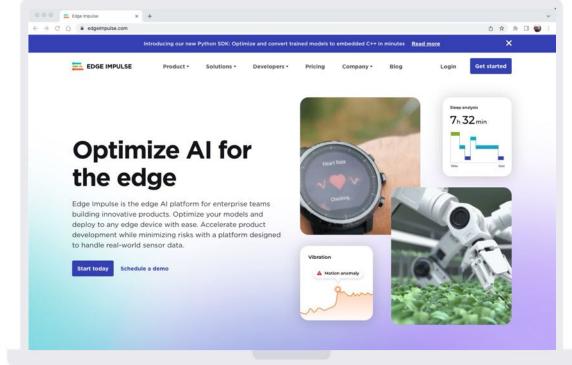






El Studio - Embedded ML platform ("AutoML")





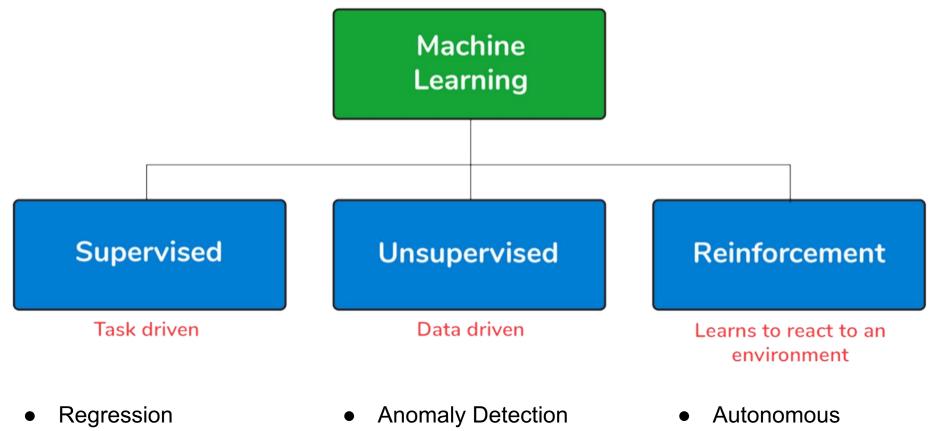
Learn more at http://edgeimpulse.com



8 Copyright © EdgeImpulse Inc.

2023 SciTinyML Workshop - Edge Impulse Overview by Shawn Hymel

TinyML Applications Examples



- Classification
- **Object Detection** ${\color{black}\bullet}$

Navigation

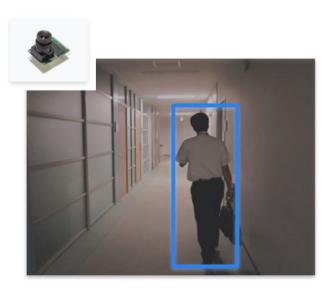
Sound

Vibration

Vision





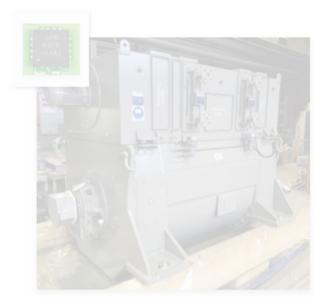


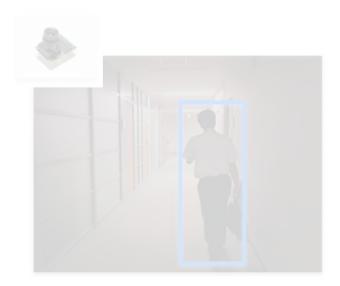
Sound

Vibration

Vision





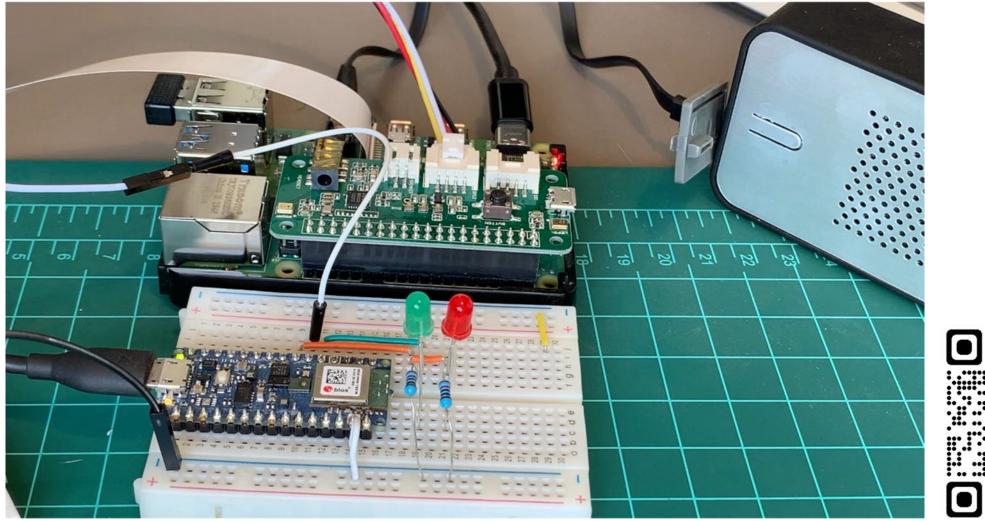


Personal Assistant

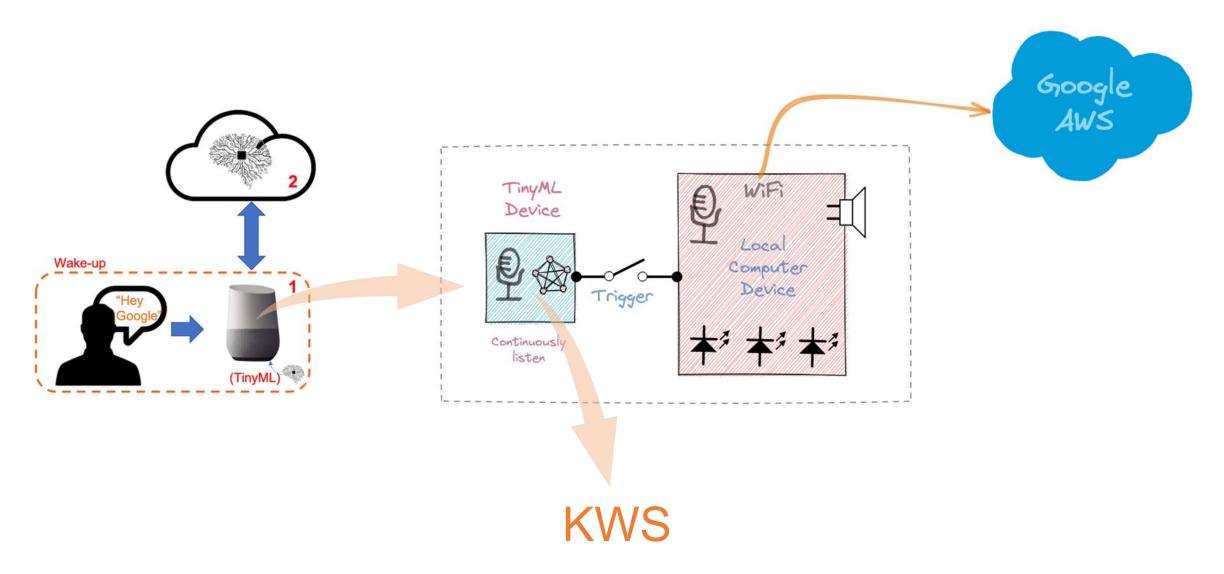




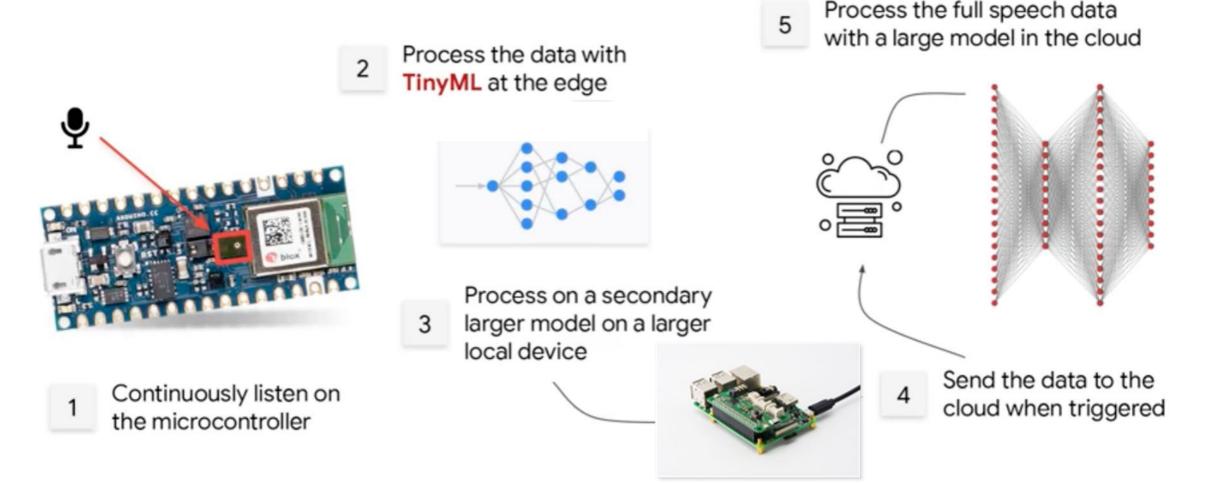
Personal Assistant



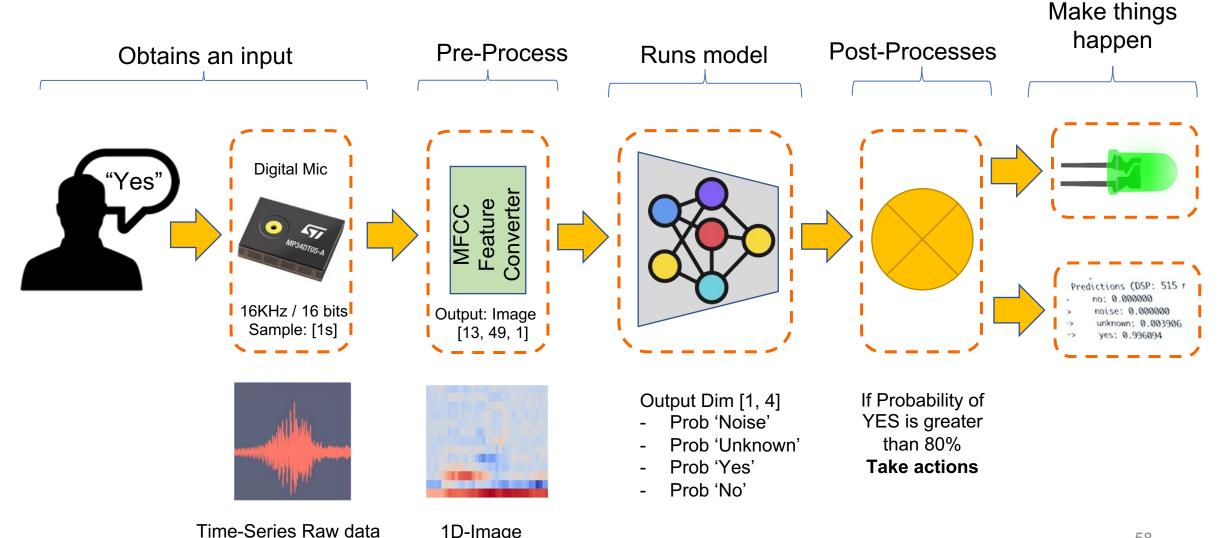
Personal Assistant

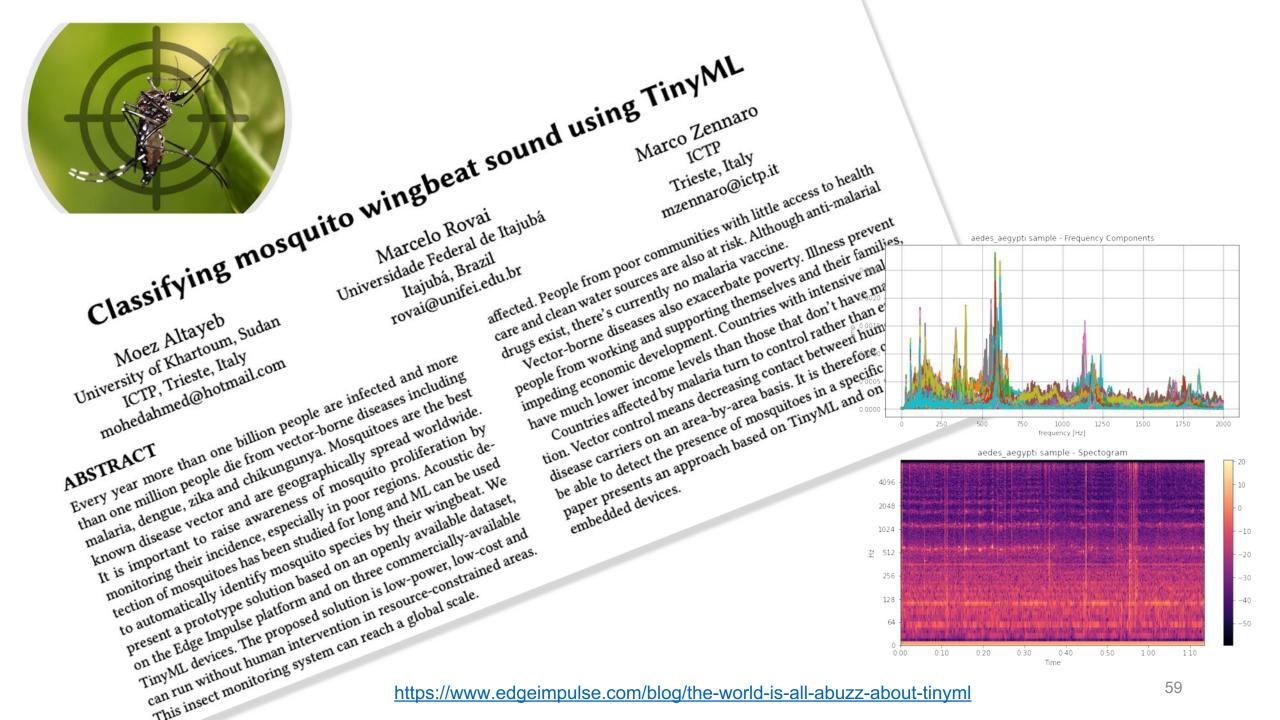


"Cascade" Detection: multi-stage model

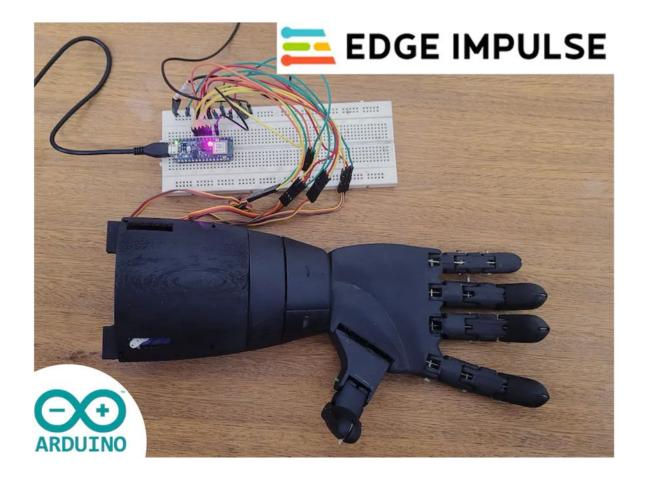


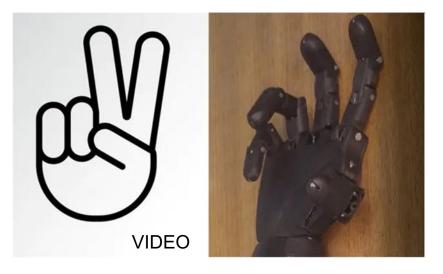
KeyWord Spotting (KWS) - Inference





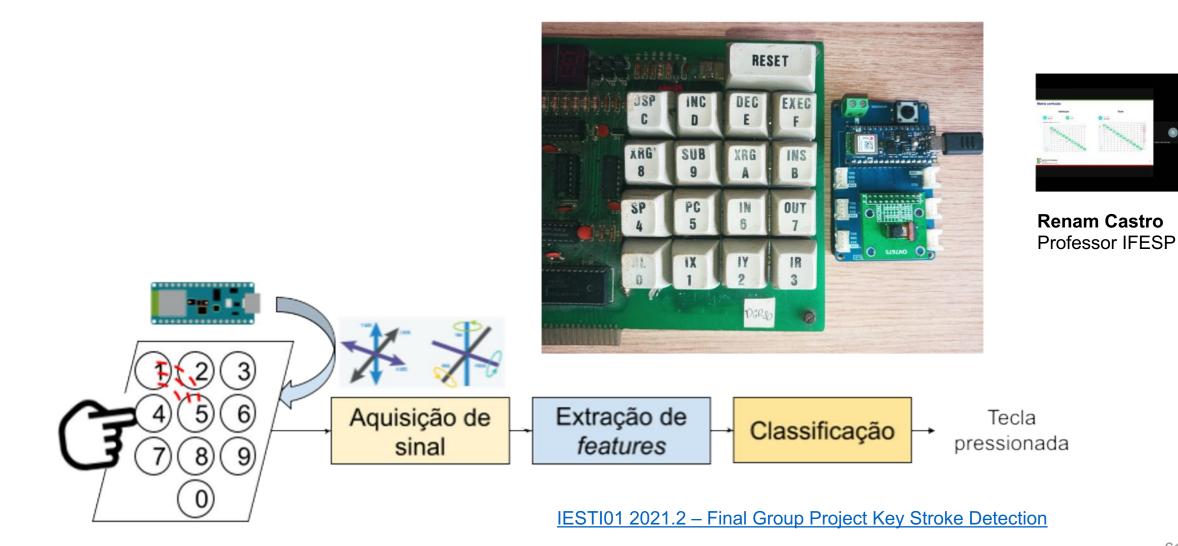
Bionic Hand Voice Commands Module





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

Keystroke Sound Detection



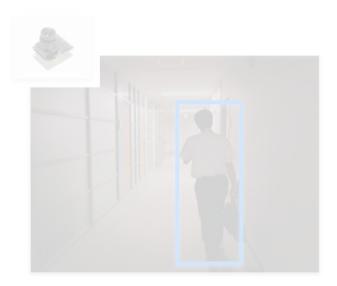
Sound

Vibration

Vision





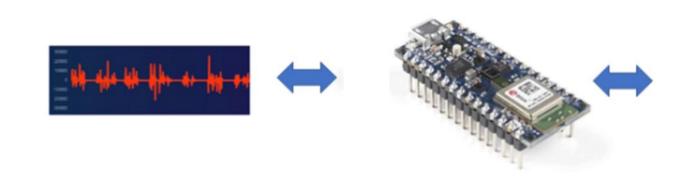


Industrial – Anomaly Detection



Machine Learning





IESTI01 2021.2 - Final Group Project: Bearing Failure Detection

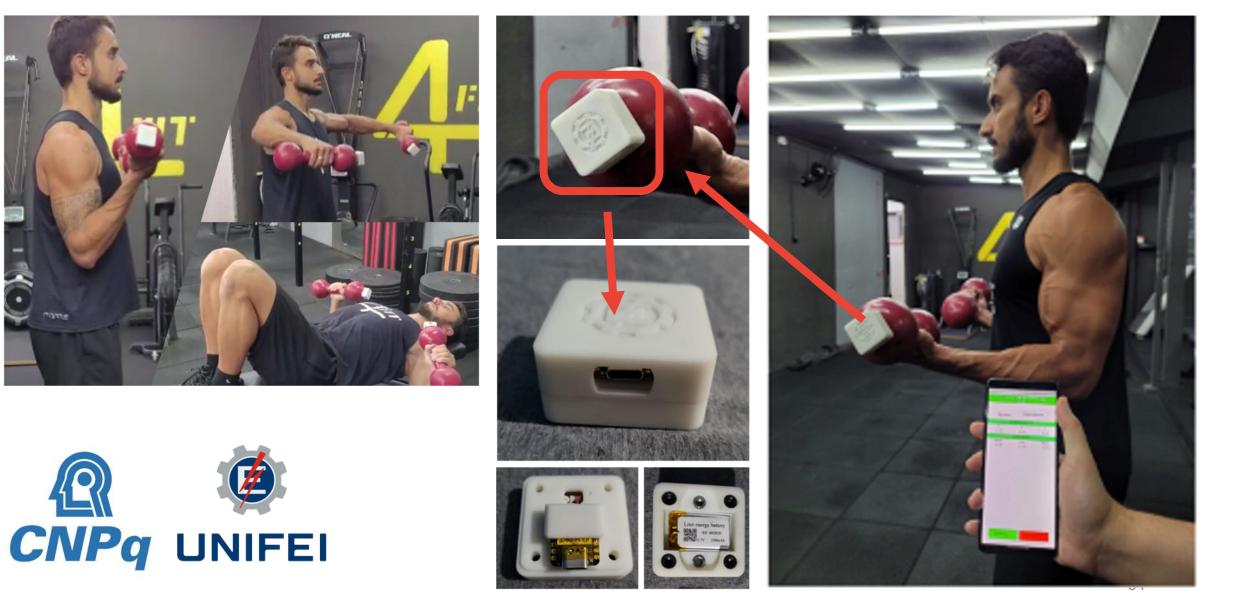


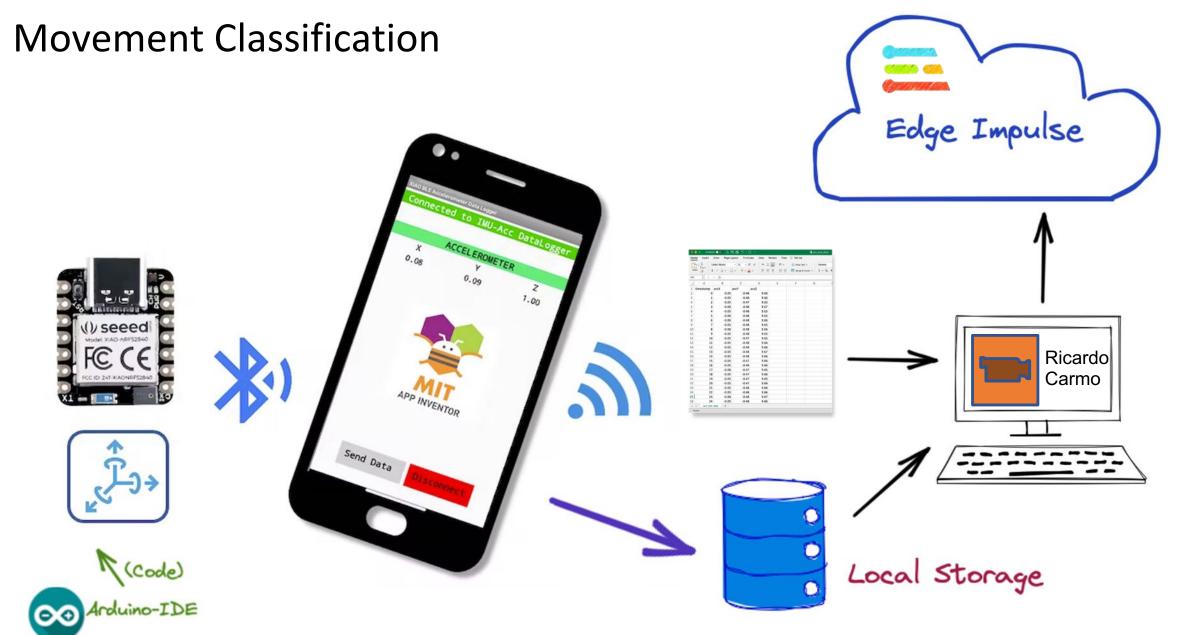






Movement Classification





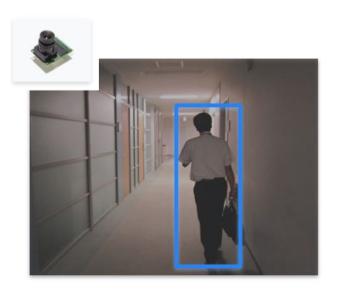
Sound

Vibration

Vision







Computer Vision Main Types

Image Classification (Multi-Class Classification)



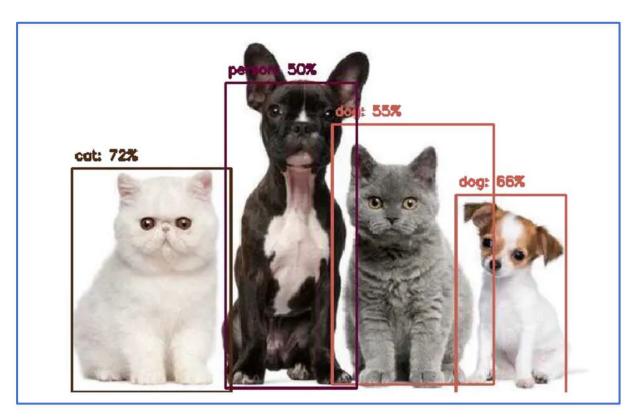
Cat: 70%

1.10

Dog: 80%

Object Detection

Multi-Label Classification + Object Localization



Computer Vision Main Types

Image Classification (Multi-Class Classification)

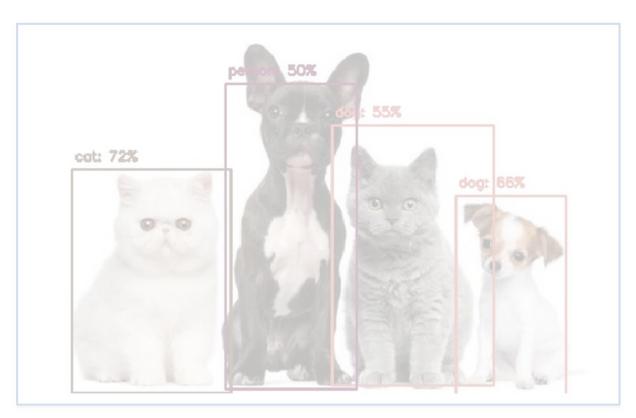


Cat: 70%

Dog: 80%

Object Detection

Multi-Label Classification + Object Localization



Forest Fire Detection



OV7675 0

IESTI01 - Forest Fire Detection – Proof of Concept

TinyML Aerial Forest Fire Detection

Coffee Disease Classification







João Vitor Yukio Bordin Yamashita Graduando em Engenharia Eletrônica pela UNIFEI

Computer Vision Main Types

Image Classification (Multi-Class Classification)

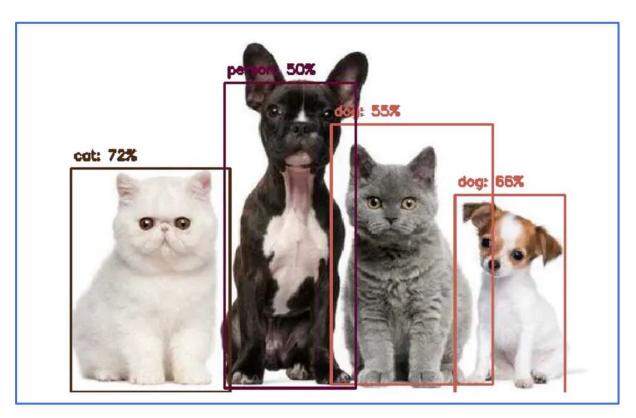


Cat: 70%

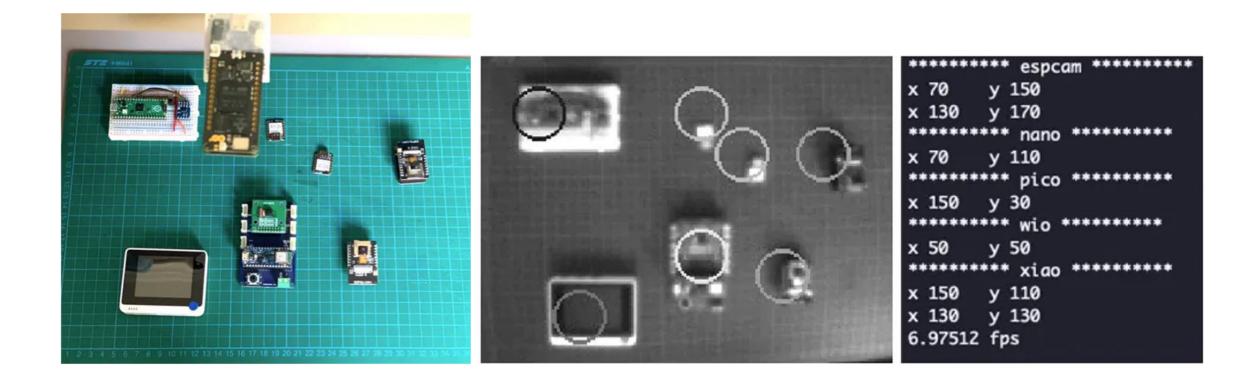
Dog: 80%

Object Detection

Multi-Label Classification + Object Localization

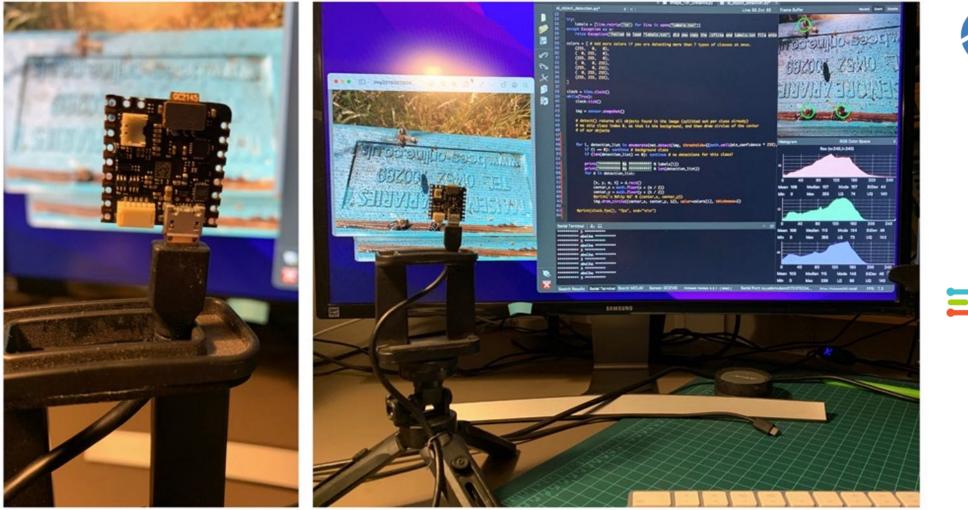


Detecting Objects using TinyML (FOMO)



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

Detecting Objects using TinyML (FOMO)





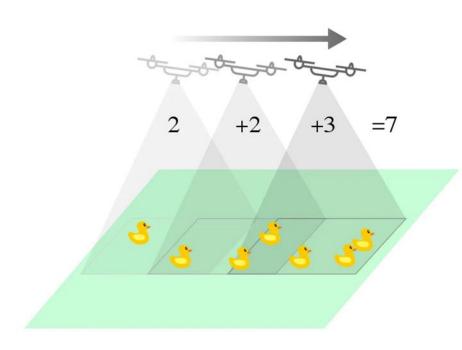


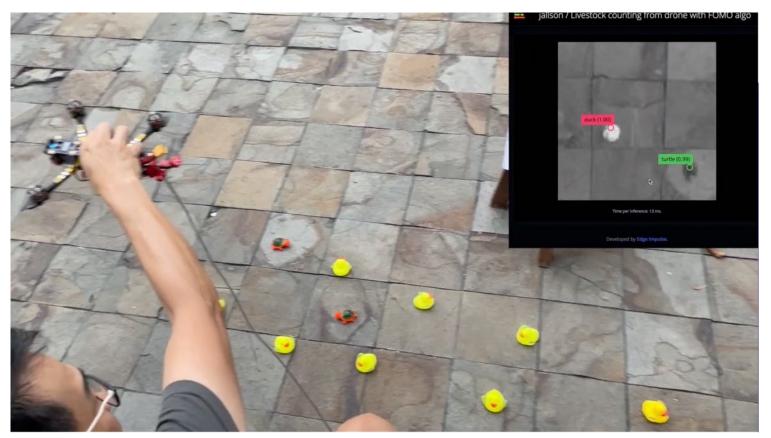
MicroPython





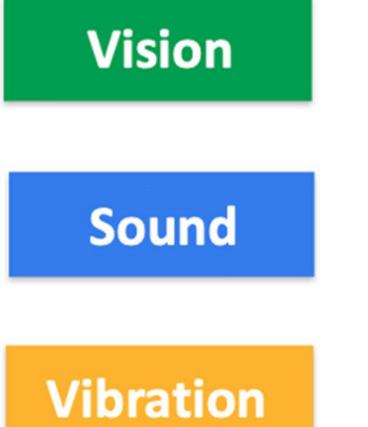
Livestock / Wildlife Counting from Drone with FOMO





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

Other TinyML / MCUs Project Examples



- Image Classification with ESP32-CAM
- Image Classification with Portenta H7
- Object Detection with Portenta H7



[Doc]

[Doc]

[Doc]

[Doc]

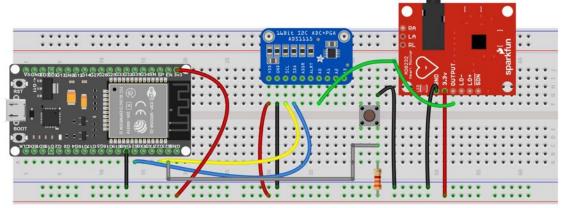
- Listening Temperature with Nano 33
- COPD Detection with Nano 33
- Sound Classification with XIAO BLE Sense [Doc]
- Motion Recognition with RPi Pico
- Gesture Recognition with Wio Terminal
- Anomaly Detection with XIAO BLE Sense [Doc]

Other Sensors / MCUs / Models Examples

AD8232 - Single Lead Heart Rate Monitor



<u>Atrial Fibrillation Detection on ECG using TinyML</u> <u>Silva et al. UNIFEI 2021</u>



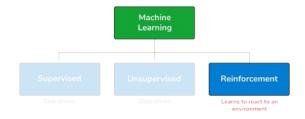
fritzing





Guilherme Silva Engenheiro - UNIFEI

Reinforcement on TinyML



Deep Reinforcement Learning for Autonomous Source Seeking on a Nano Drone

Bardienus P. Duisterhof^{1,3} Srivatsan Krishnan¹ Jonathan J. Cruz¹ Colby R. Banbury¹ William Fu¹

Aleksandra Faust² Guido C. H. E. de Croon³ Vijay Janapa Reddi^{1,4}

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



https://youtu.be/wmVKbX7MOnU

TinyML Academic Network

Widening access to applied machine learning by establishing best practices in education.



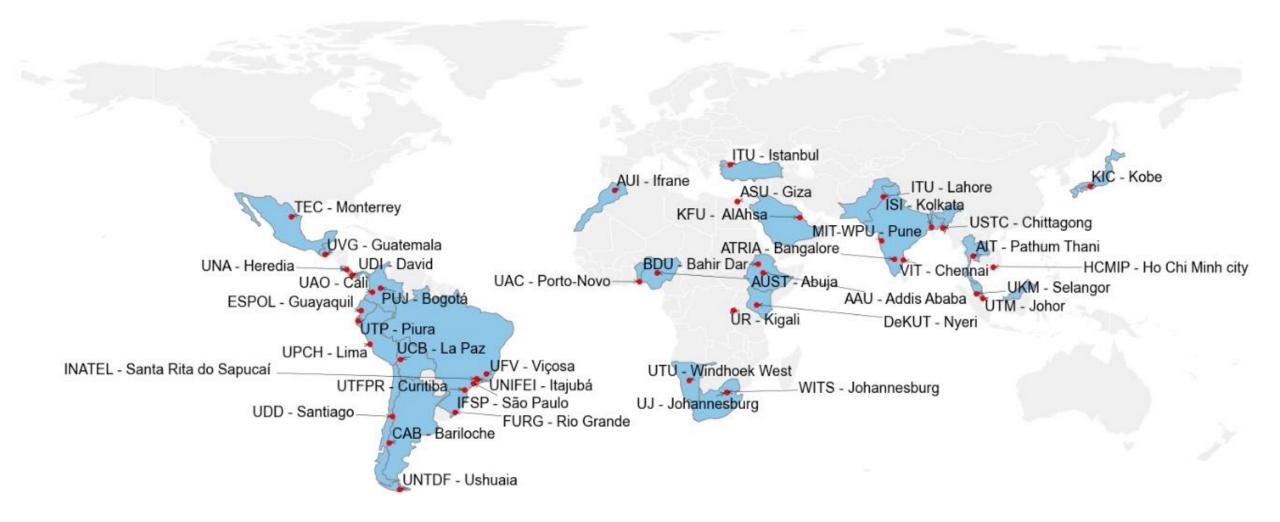
Harvard John A. Paulson School of Engineering and Applied Sciences



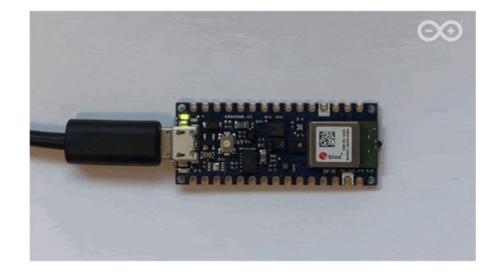




TinyML4D Academic Network - March 2023



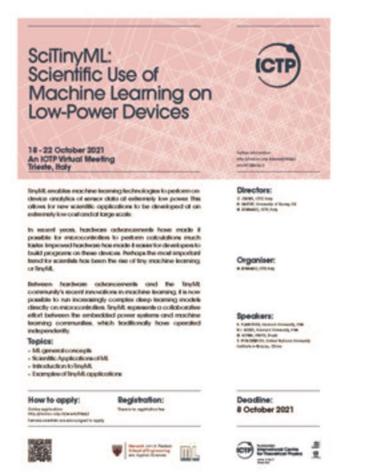


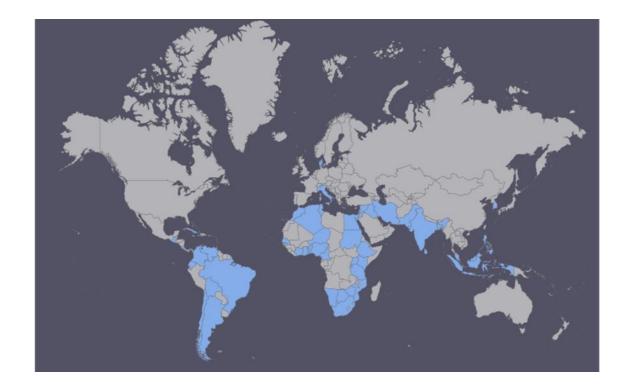


tinymledu.org

🗢 🗢 💼 Welcome to the Tiny Machine 🗆 🗙 🕂				v :			
← → C ☆ 🌲 tinyml.seas.harvard.edu			🖞 🖈 🗟 🕿	🛢 🕈 🕼 🚳 🌇 e 🔳 🐨 🛪 🗊 🚳 🔅			
TinyML Harvard Funding MLC Meta	😳 CS141 💧 CS141 🚼 Energy Systems C 🔁 Nora	a 🔯 ChatGPT 🛆 DESC 🛆 DOE 2023 😵 DOE_SC_FOA_00		Cther Bookmarks			
	TinyMLedu Home	Courses & Materials 4D Network 5	Show & Tell SciTinyML Research				
Welcome to the Tiny Machine Learning Open Education Initiative (TinyMLedu)							
	Take a Free Course or Teach You	r Own Explore our 4D Academic Netwo	rk Attend our SciTinyML Workshop				
	View our Research Projects	Learn More About Us					
	اـــــــــــــــــــــــــــــــــــــ						
If you want to be more involved with our effort to help improve access to TinyML educational materials and hardware resources worldwide reach out to us at <u>edu@tinyML.org</u> !							
	Harvard John A. Pau School of Engineer and Applied Science	ing AZZ	Google				

2021 activities





210 participants from 48 countries

2022 activities



187 participants29 African countries

100 participants8 Asian countries

183 participants 17 LatAm countries

April 2023 ICTP virtual workshop

Day

Date

Topics

Workshop on Scientific Use of Machine Learning on Low-Power Devices: (CTP) Applications and Advanced Topics

17 - 21 April 2023 An ICTP Virtual Meeting Trieste, Italy

TinyML is a subfield of Machine Learning focused on developing models that can be executed on small, realtime, low-power, and low-cost embedded devices. This allows for new scientific applications to be developed at an extremely low cost and at large scale

Topics:

in DryML

Description:

InvML represents a collaborative effort between the embedded power systems and Machine learning communities, which dollaroally have operated independently. Getting started with the linvML kit traditionally have operated independently. Examples of TinyML Applications

TervML hos a significant role to play in achieving the SDGs and facilitating scientific research in oreas such as environmental monitoring, physics of complex systems and environmental energy management.

The TinyAL process starts with collecting data torm toil devices, then training the collected dataset to estratch knowledge patients these patients are then packaged into a TinyAL, model that considers the target microprocessor's timbed resources such on memory, processing power, and energy.

Through hands-on examples, this workshop will focus on both introductory and advanced topics in TinyML to pave the way to the development of real-world polications

How to apply: Online application: http://indico.iotp.it/event/10146/

UNIFEI

Deadline:

immite scientists are encouraned to contin

(B)

BARNARD

Harvard John A. Paulson School of Engineering and Applied Sciences .

Registration:

There is no registration fee.

Recent Research and Advanced Topics



(6)

http://indica.ictp.

mr3833Rictp.B

Directors:

B. PLANCHER, Barnard College, USA V. J. BEDDI, Harvard University, USA M. BOVAL, Redenal University of Itajubó, Brazil

Local Organiser:

M. ZENNARO, ICTP, Ibaly

SciTinyMI

Scient Learni (CTP)

TinyMI

Home

Sched Call for Team

Update

by @pl

tific Use of Machine ing on Low-Power Devices	Day 1	Monday	Introduction to (tiny)ML 10:00 AM Workshop Opening and Schedule 10:30 AM Keynote 11:15 AM Introduction to Machine Learning 12:15 PM Introduction to Embedded ML 12:55 PM Day Closing	Marco Zennaro of ICTP Diego Mendez Chaves of Pontificia Universidad Javeriana Robert Thas John of Versus
ALedu 🖸	Day 2	Tuesday	Hands-on Introduction to TinyML 10:00 AM Day Opening 10:05 AM Edge Impulse Overview and New Features 10:40 AM Hands-on Motion Classification and Anomaly Detection 12:10 PM AI Ethics: Avoiding Bias 12:55 PM Day Closing	Marco Zennaro of ICTP Shawn Hymel of Edge Impulse José Antonio Bagur Nájera of Universidad del Valle de Guatemala Viola Schiaffonati and Manuel Roveri of Politecnico di Milano
	Day 3	Wednesday	Expanding Your Options and Devices 10:00 AM Day Opening 10:05 AM Leveraging Other Microcontrollers and Sensors 11:00 AM WebUSB and FOMO 11:30 AM Adding IoT to a Project with Blues Wireless 11:55 AM Industry 5.0 with Jetson Nano 12:25 PM MLOps: Scaling Deployments 12:55 PM Day Closing	Marceo Zennaro of ICTP Marcelo Rovai of Federal University of Itajuba - UNIFEI Jeremy Ellis of School District 75 Mission Peter Ing of TFG (The Foschini Group) Marcelo Pias of Federal University of Rio Grande FURG Colby Banbury of Harvard University
	Day 4	Thursday	TinyML Show and Tell 10:00 AM Day Opening 10:05 AM Selected Show and Tell Talks 12:55 PM Day Closing	Brian Plancher of Barnard College, Columbia University
	Day 5	Friday	Advanced Scientific TinyML 10:00 AM Day Opening 10:05 AM Scientific Applications of TinyML 1 10:50 AM TinyML and Sustainability 11:20 AM TinyML and Robotics 11:50 AM Scientific Applications of TinyML 2 12:55 PM Workshop Closing and Future Events	Marco Zennaro of ICTP Matthew Stewart of Harvard University Bardienus Duisterhof of Carnegie Mellon University - CMU

Speakers and Materials

Local/Relevant Applications

Timothy Kudzanayi Kuhamba

Thanx Marisa^{a*}, Munyaradzi Munochiveyi^{a*}, Wadzanai Julius Zondai^{a*}, Ramson Munyaradzi Nyamukondiwa^a, Isheanesu Newengo^b

Case Study Zimbabwe

A DEEP LEARNING BASED APPROACH FOR FOOT AND MOUTH DISEASE DETECTION



Local/Relevant Applications

VEGETABLE DISEASE AND INSECT PEST RECOGNITION BASED ON TINYML: Cotton Case in Benin

James O. ADEOLA IMSP - UAC

Dr Marco Zenaro ICTP Italy

Dr Jules DEGILA IMSP Benin



Local/Relevant Applications

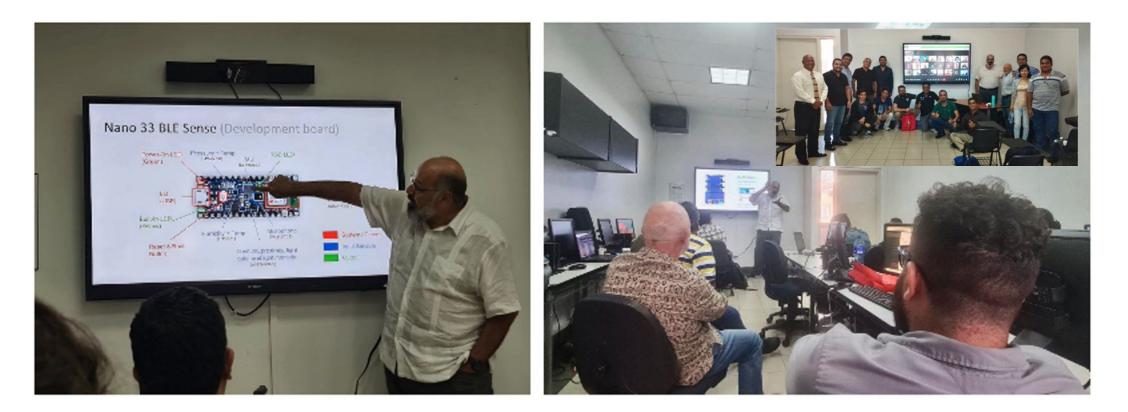


Outreach activities



Solomon from Ethiopia

Marcelo Rovai at WALC 22 - Udl, Panama



- 5 days Workshop
- 46 remote and 16 in site participants from Latin America

TinyML Academic Network @ UN 2022

Science-Policy Brief for the Multistakeholder Forum on Science, Technology and Innovation for the SDGs, May 2022

TinyML: Applied AI for Development

Marco Zennaro (ICTP/UNESCO), Brian Plancher (Harvard University), Vijay Janapa Reddi (Harvard University)

Abstract

Artificial intelligence (AI) will likely be an instrumental part of progress towards the United Nations' Sustainable Development Goals (SDGs). However, its adoption and impact are limited by the immense power consumption, strong connectivity requirements and high costs of cloud-based deployments. TinyML is a new technology that allows machine learning (ML) models to run on low-cost, low-power microcontrollers, circumventing many of these issues. We believe that TinyML has a significant role to play in achieving the SDGs and facilitating scientific research in areas such as environmental monitoring, physics of complex systems and energy management. To broaden access and participation and increase the impact of this new technology, we present an initiative that is creating and supporting a global network of academic institutions working on TinyML in developing countries. We suggest the development of additional open educational resources, South–South academic collaboration and pilot projects of at-scale TinyML solutions aimed at addressing the SDGs.

Challenges with Machine Learning in Developing Countries

Machine learning has a huge potential to tackle societal issues in diverse fields that include agriculture, conservation and healthcare. A recent study [1] main energy consumer component of an embedded system.

3. **Privacy:** Applications that send data from the point of collection to the cloud may leak private information as data must be transmitted over the internet.

TinyML Academic Network @ UN 2023

Bridging the Digital Divide: the Promising Impact of TinyML for Developing Countries

Marco Zennaro (ICTP/UNESCO), Brian Plancher (Barnard College, Columbia University), Vijay Janapa Reddi (Harvard University)

Abstract

The rise of TinyML has opened up new opportunities for the development of smart, low-power devices in resource-constrained environments. This technology has particular relevance for developing countries, where access to energy and computing resources is often limited. In light of this, a network of 40 universities has been established over the past two years with the goal of promoting the use of TinyML in developing regions. The members of this network have taught courses at their home institutions and have completed their first research projects covering topics ranging from the diagnosis of respiratory diseases in Rwanda to assistive technology development in Brazil, bee population monitoring in Kenya and estimating the lifespan of the date palm fruit in Saudi Arabia. These initial projects demonstrate the potential

Papers published

Mihigo, Irene Niyonambaza, et al. "On-Device IoT-Based Predictive Maintenance Analytics Model: Comparing TinyLSTM and TinyModel from Edge Impulse." Sensors 22.14 (2022): 5174.

Altayeb, Moez, Marco Zennaro, and Marcelo Rovai. "Classifying mosquito wingbeat sound using TinyML." Proceedings of the 2022 ACM Conference on Information Technology for Social Good. 2022.

Bamoumen, Hatim, et al. "How TinyML Can be Leveraged to Solve Environmental Problems: A Survey." 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT). IEEE, 2022.

Avellaneda, Diego, Diego Mendez, and Giancarlo Fortino. "A TinyML Deep Learning Approach for Indoor Tracking of Assets." Sensors 23.3 (2023): 1542.

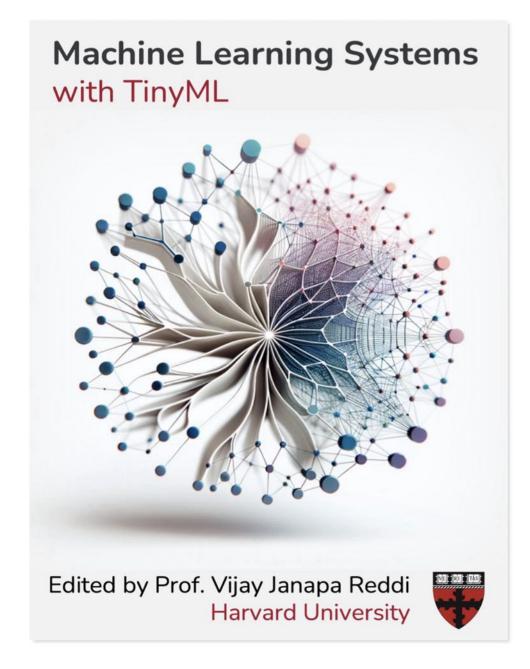
Papers published

Avellenada, Diego, Diego Mendez, and Giancarlo Fortino. "BLE-based Indoor Positioning Platform Utilizing Edge Tiny Machine Learning." 2022 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech). IEEE, 2022.

Plancher, Brian, and Vijay Janapa Reddi. "TinyMLedu: The tiny machine learning open education initiative." Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 2. 2022.

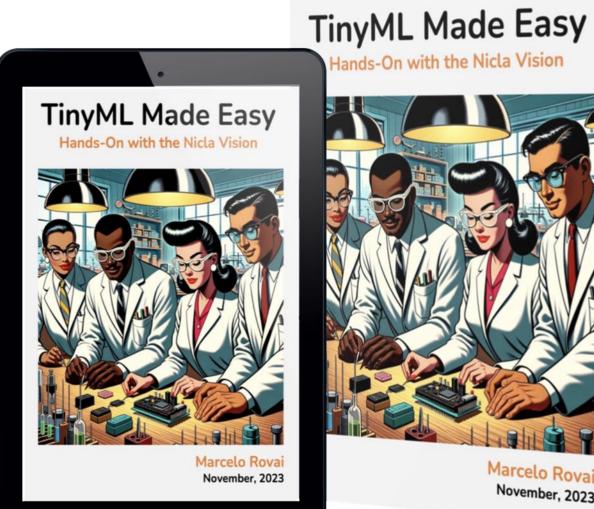
G. Silva, M.D. Lima, J.A.F. Filho and M.J. Rovai " "Atrial Fibrillation and Sinus Rhythm detection using TinyML (Embedded Machine Learning). "IX Latin American Congress on Biomedical Engineering" and "XXVIII Brazilian Congress on Biomedical Engineering"

João Vitor Yamashita et al.. "Coffee disease classification at the edge using deep learning". Smart Agricultural Technology Volume 4, August 2023, 100183





https://github.com/harvard-edge/cs249r_book







https://github.com/Mjrovai/TinyML_Made_Easy_NiclaV_eBook

Thanks

