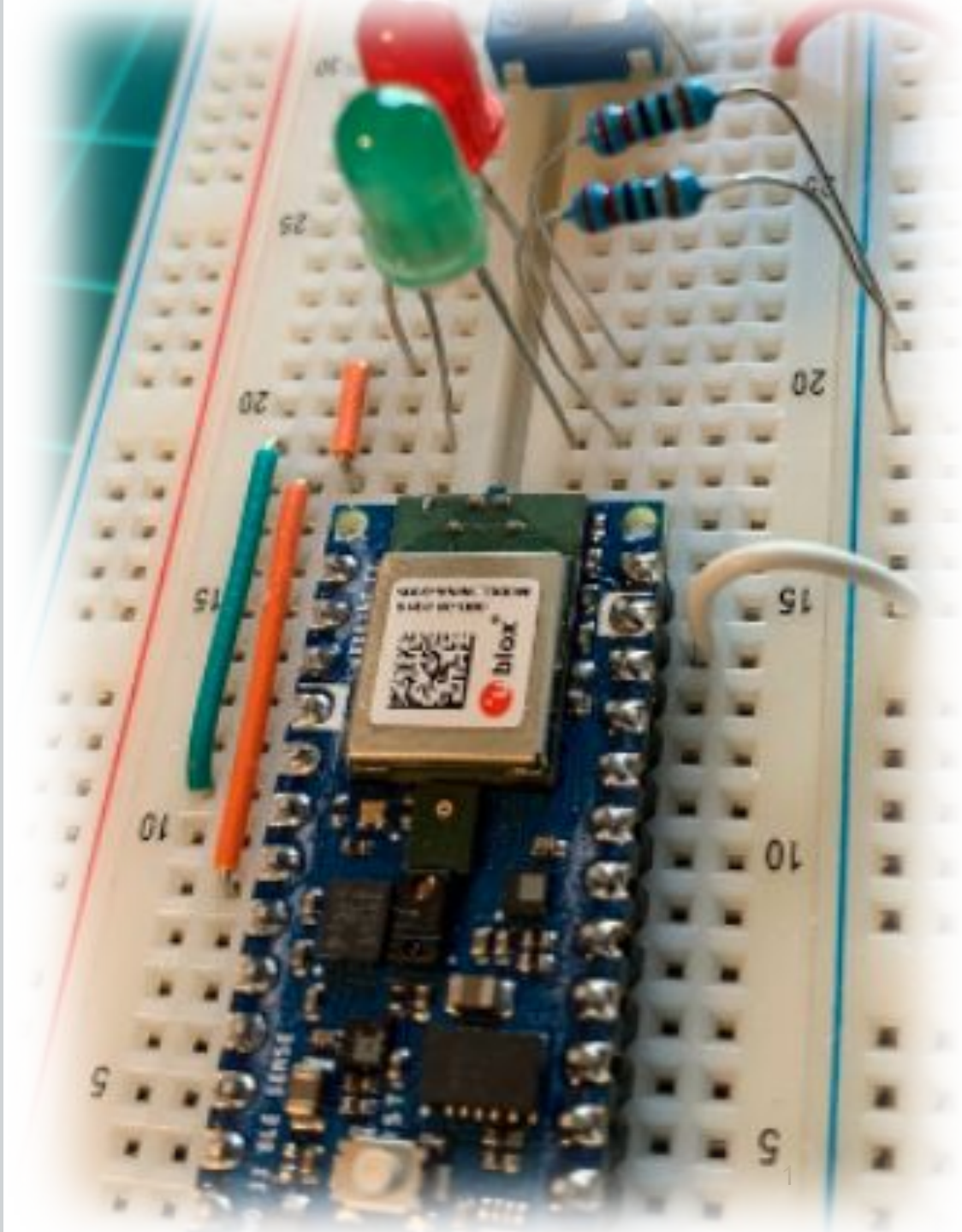
IESTI01 – TinyML

Embedded Machine Learning

1. About the Course & Syllabus



Prof. Marcelo Rovai
UNIFEI

Dear students,

Welcome to the first class of the IESTI01 (TinyML) 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 type of discipline (we are the second university worldwide to have this type of course, being Harvard School of Engineering the first) leads us to significant challenges concerning the necessary basis for the minimum understanding of the matter. Thus, it is essential to emphasize that the time we have available for the course is short. So, We must commit ourselves to do our utmost in class and with complementary activities such as readings, laboratories, and assignments. You can count on me outside the hours of the weekly classes to clarify doubts and to review the necessary concepts for a good understanding of the course.

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, where continue as 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

What is TinyML?

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

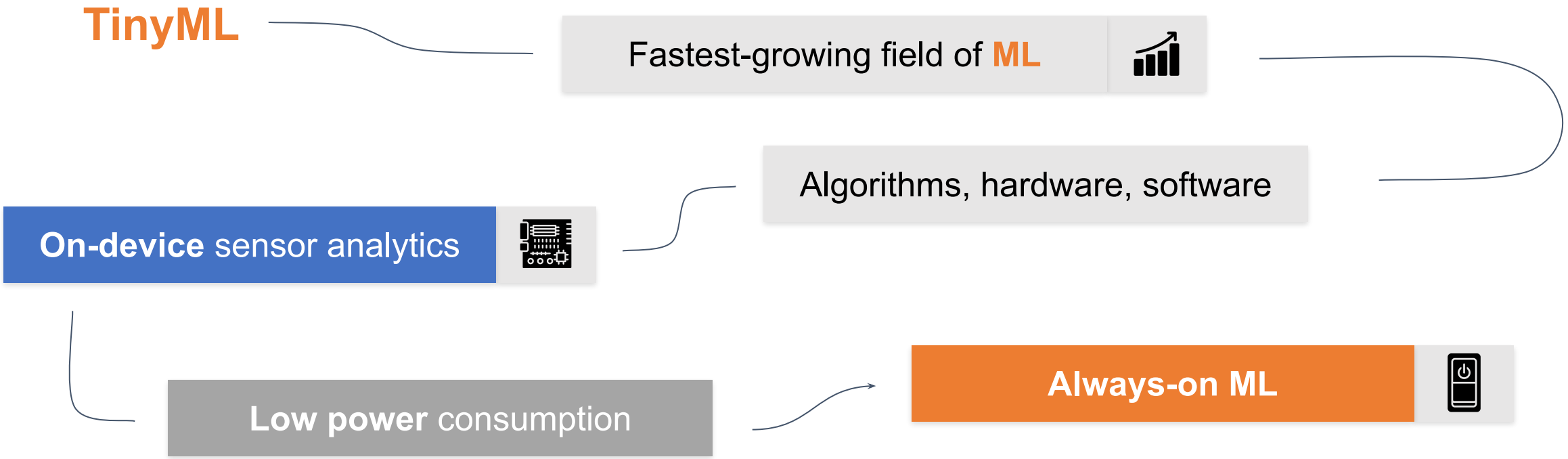
Fastest-growing field of **ML**

Algorithms, hardware, software

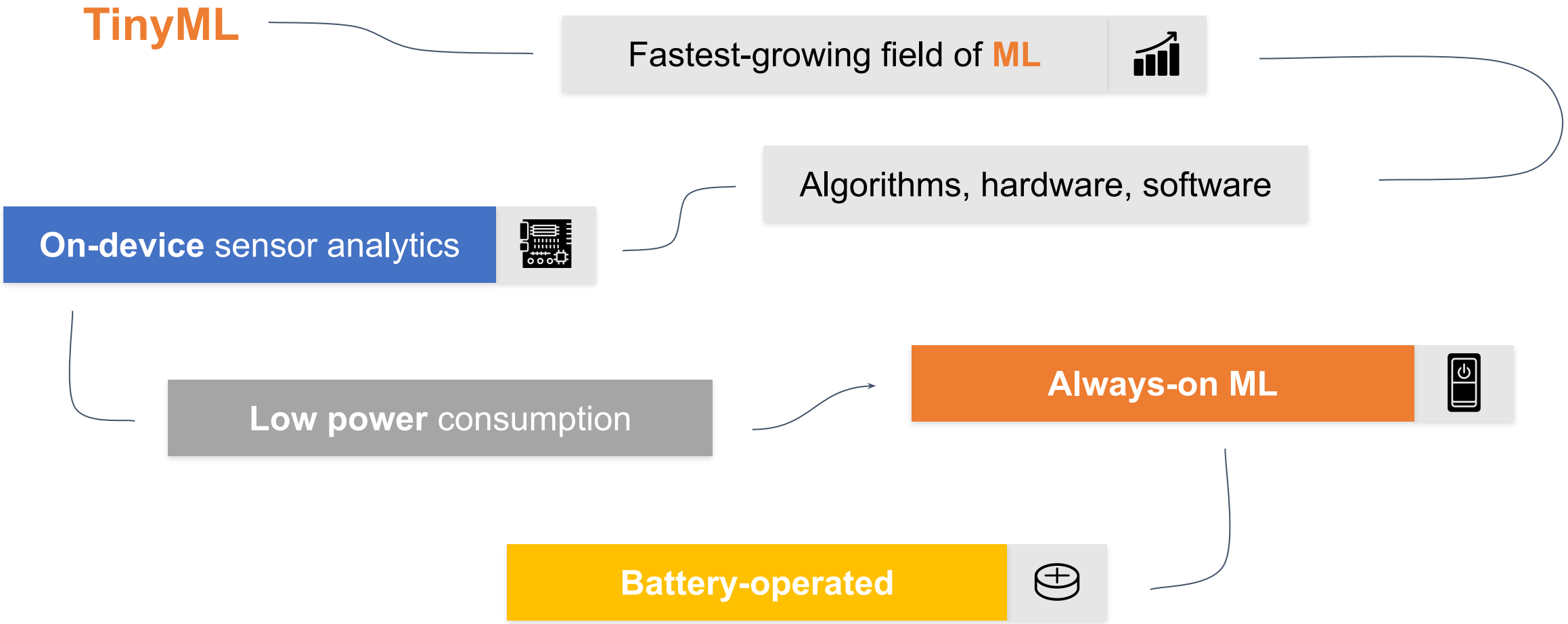
On-device sensor analytics

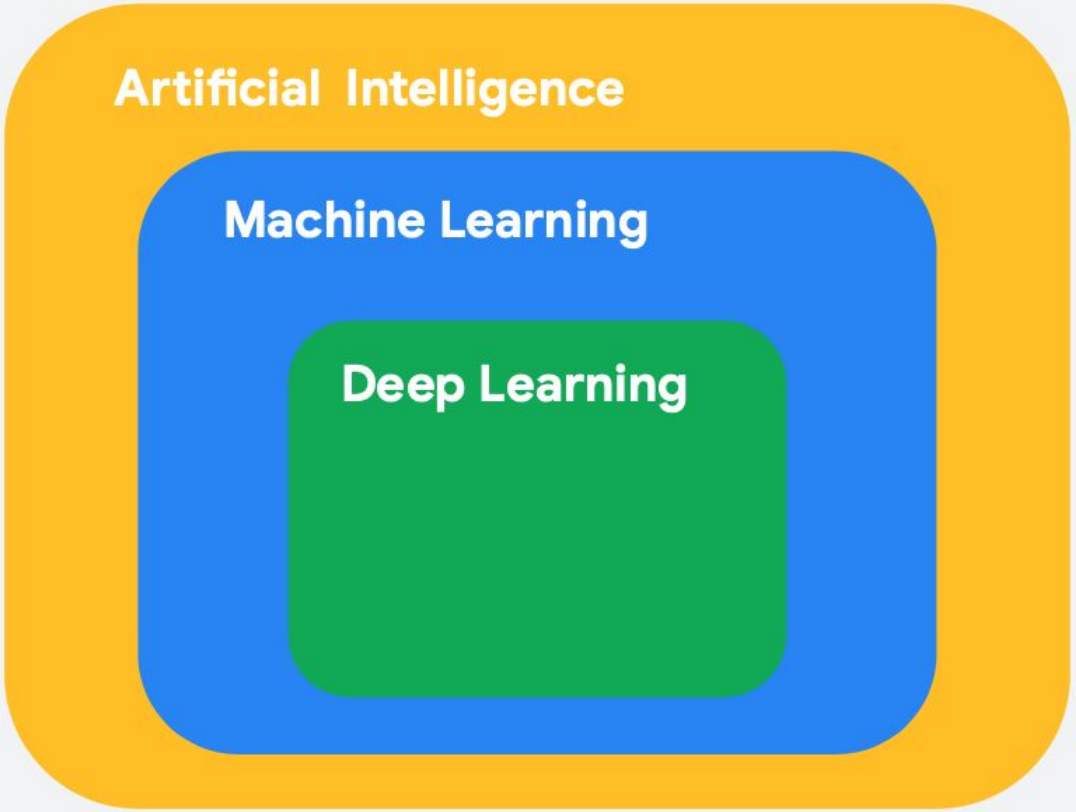
Low power consumption

What is Tiny Machine Learning (TinyML)?



What is Tiny Machine Learning (TinyML)?

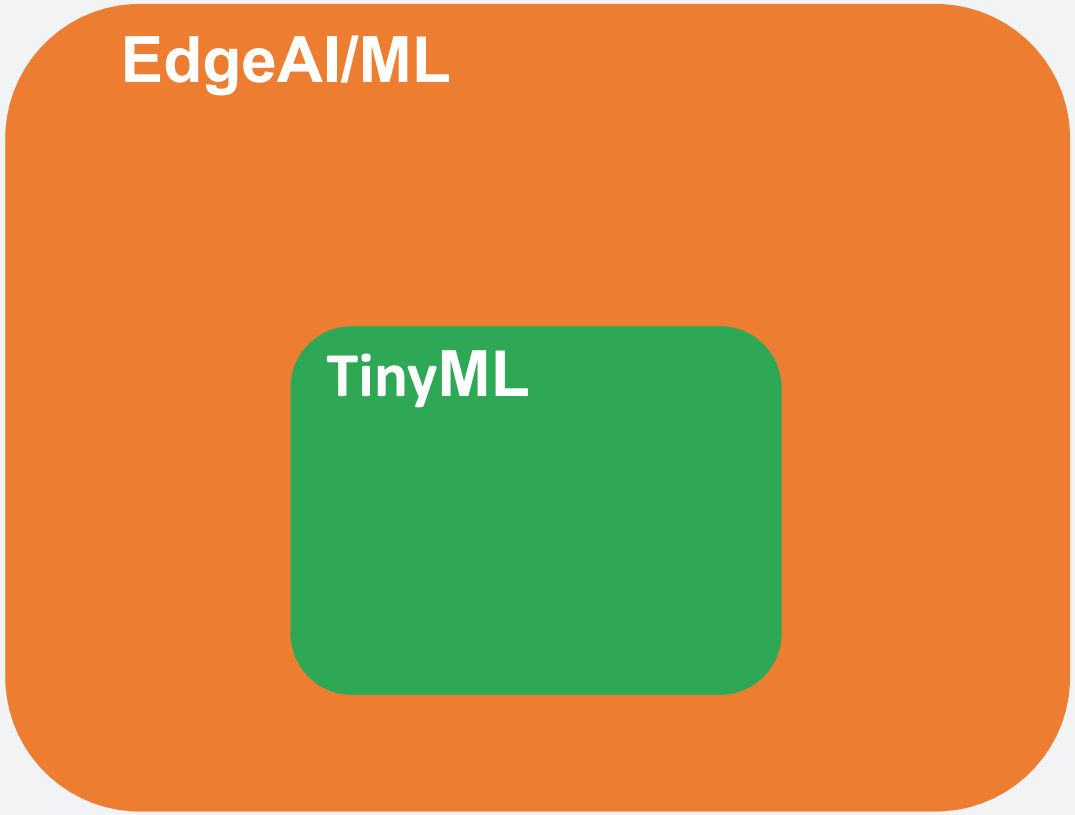




AI: Any technique that enables computers to mimic human behavior

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")

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

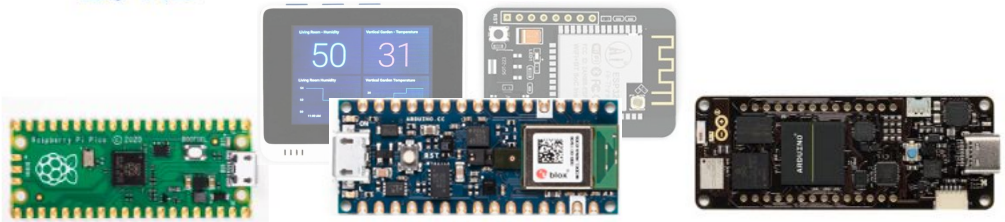
Video Classification 2 MB+

Object Detection Complex Voice Processing 1 MB+

Image Classification 250 KB+

KeyWord Spotting Audio Classification 50 KB

Anomaly Detection Sensor Classification 20 KB



Rpi-Pico (Cortex-M0+)

Arduino Nano (Cortex-M4)

Arduino Pro (Cortex-M7)

RaspberryPi SmartPhone (Cortex-A)

Jetson Nano (Cortex-A + GPU)

Source: Edge Impulse

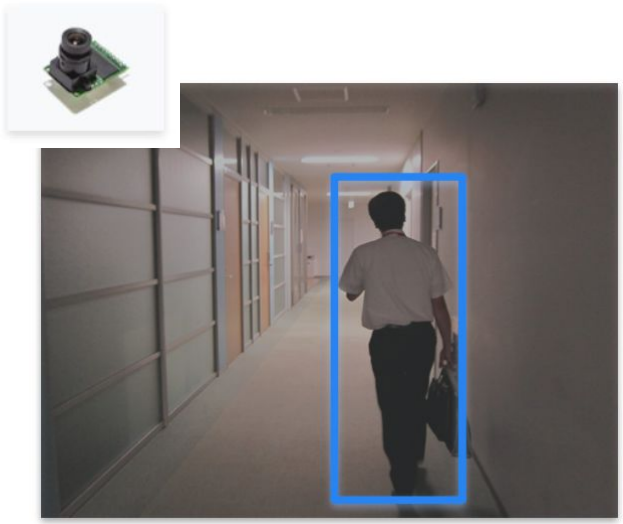
TinyML Application

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

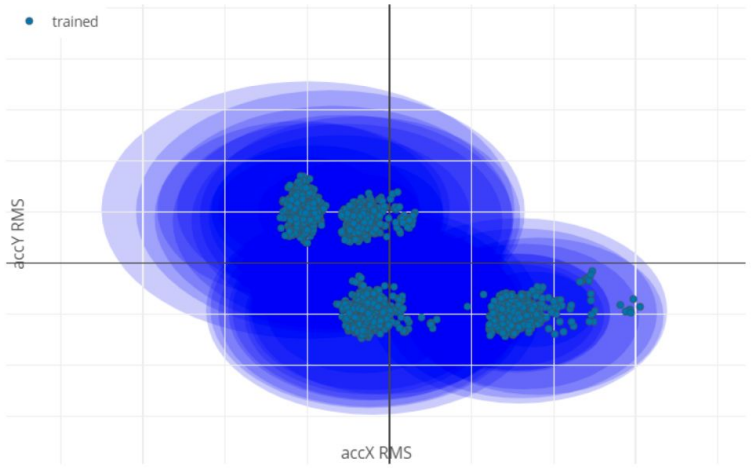
EDGE IMPULSE

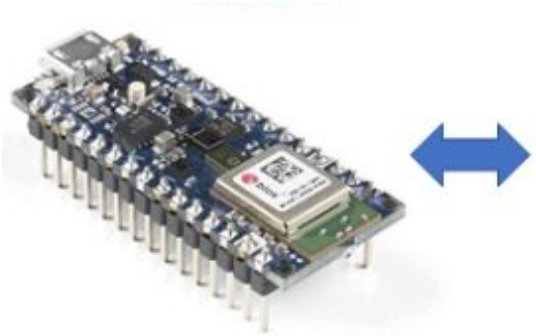
Application: Factory machinery

Ball Bearings

Accelerometer

Industry – Anomaly Detection





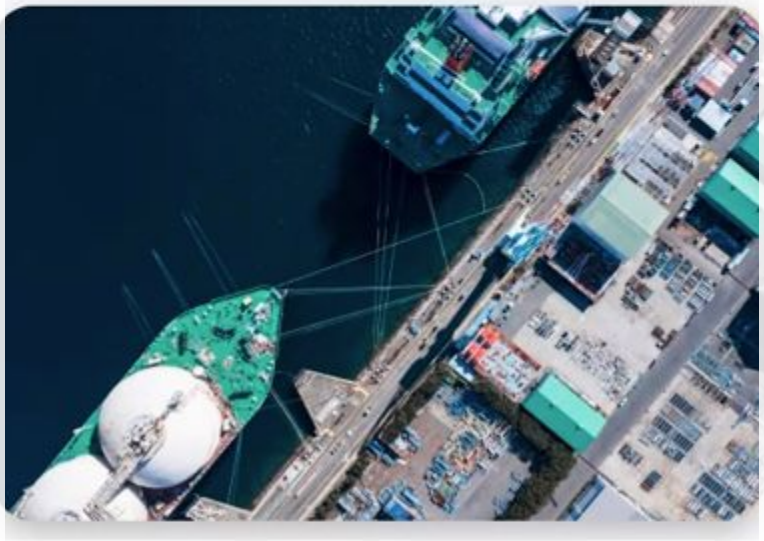
[IESTI01 2021.2 - Final Group Project: Bearing Failure Detection]

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**

EDGE IMPULSE

Human & Animal Sensing

Motion, radar, audio, PPG, ECG

- Health
- Consumer
- Industrial

Cow Monitoring

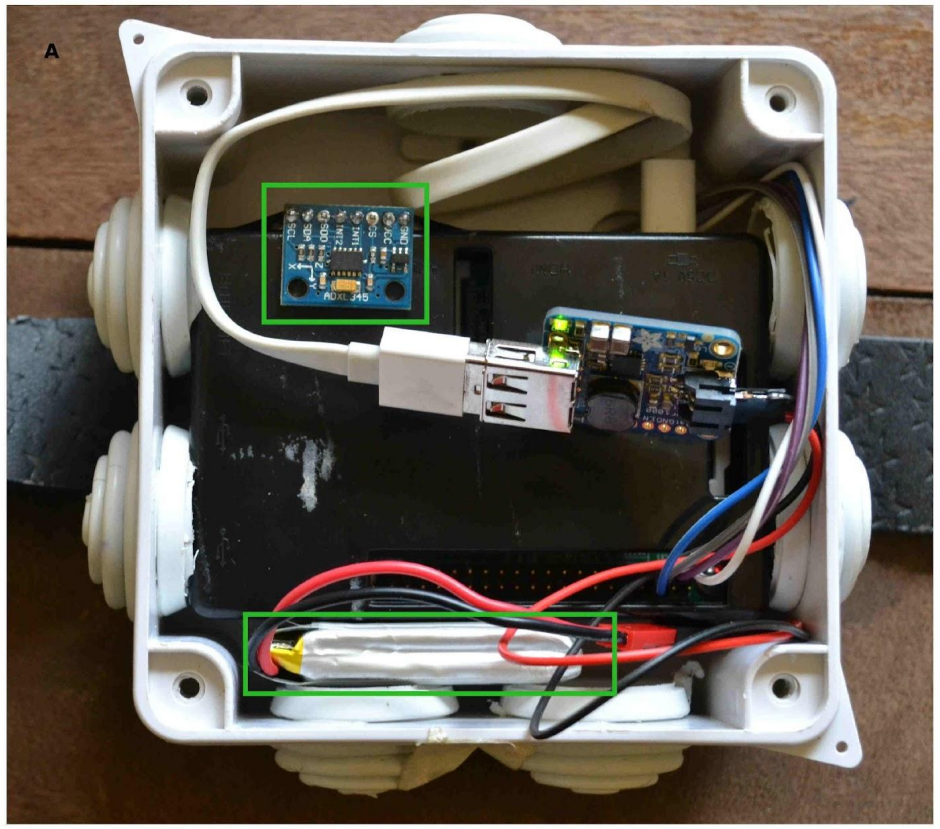
Using the Internet of Things for Agricultural Monitoring

"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

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

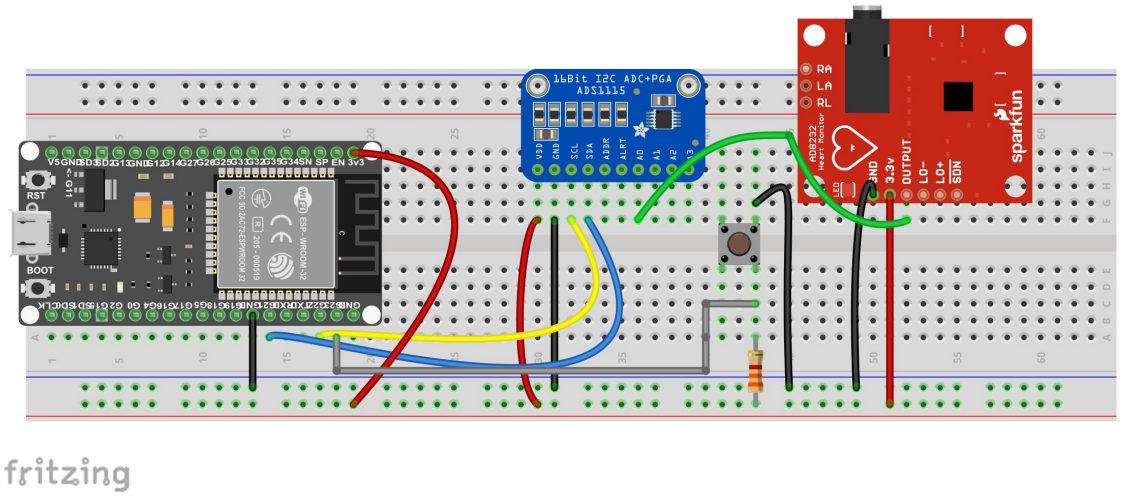
Human & Animal Sensing

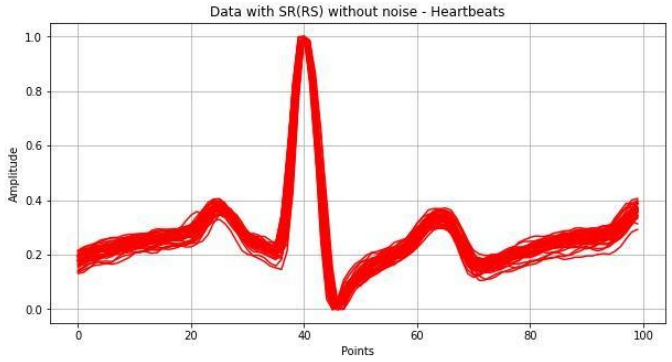
Motion, radar, audio, PPG, ECG

- **Health**
- Consumer
- Industrial

EDGE IMPULSE

Health - Human Sensing





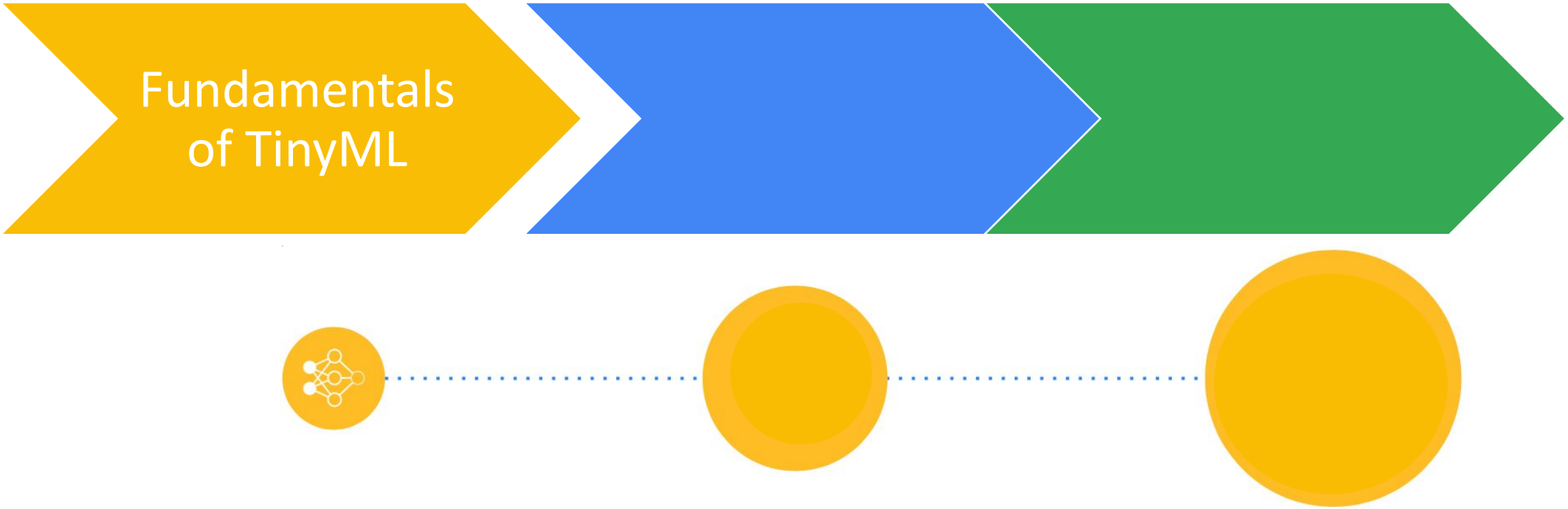
Guilherme Silva
Engenheiro - UNIFEI

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

What will We learn?

What will We learn?

Part 1

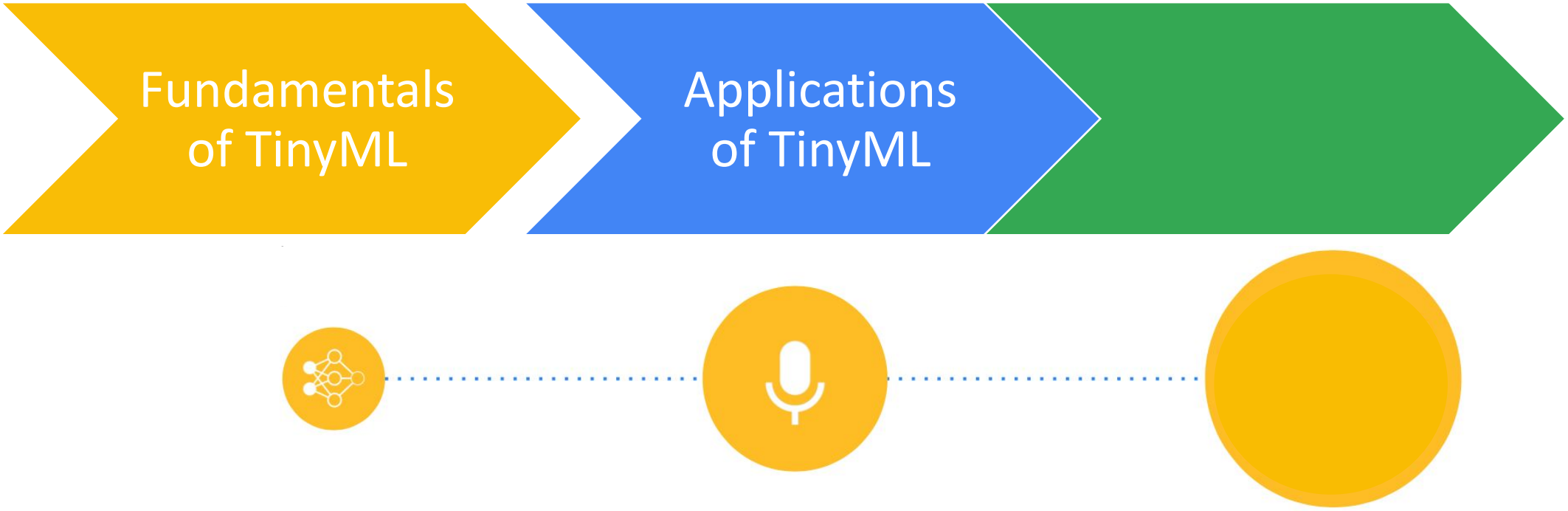


Part 1 is all about talking about what is the language of machine learning

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?

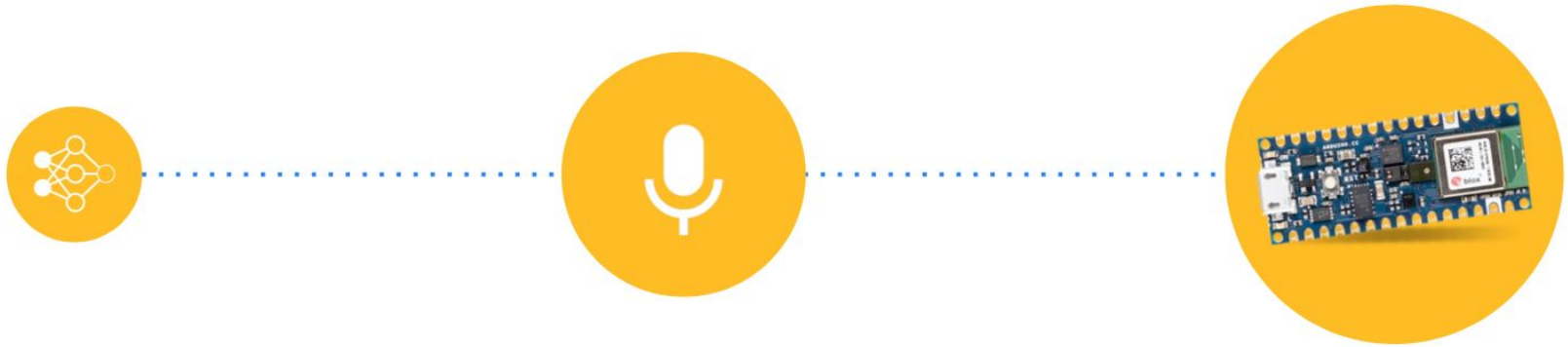
Part 1

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

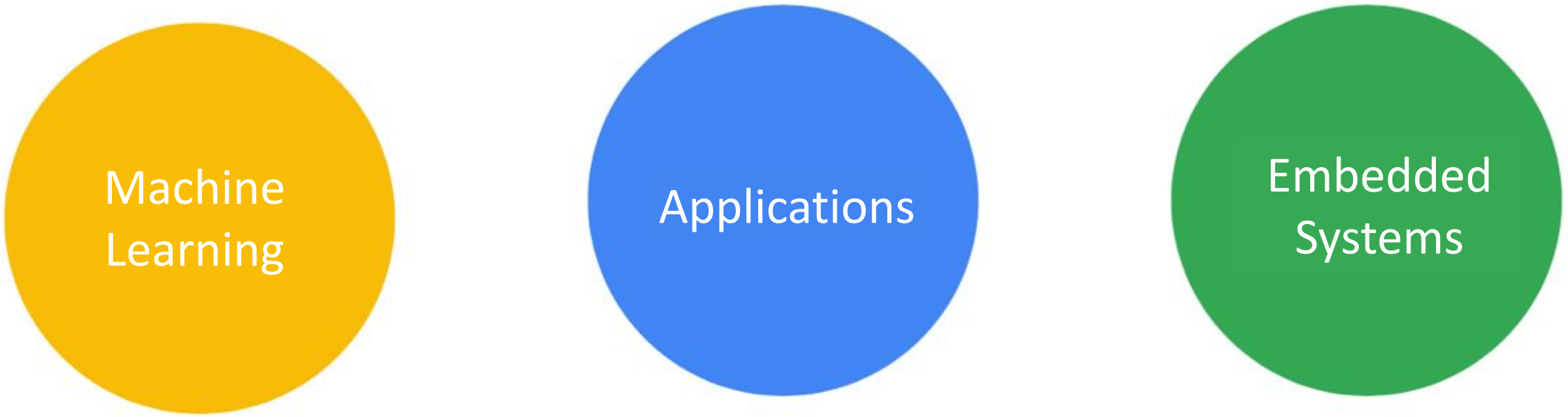
- Python
- TensorFlow
- Google Colab
  - Jupyter Notebook

- 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**.

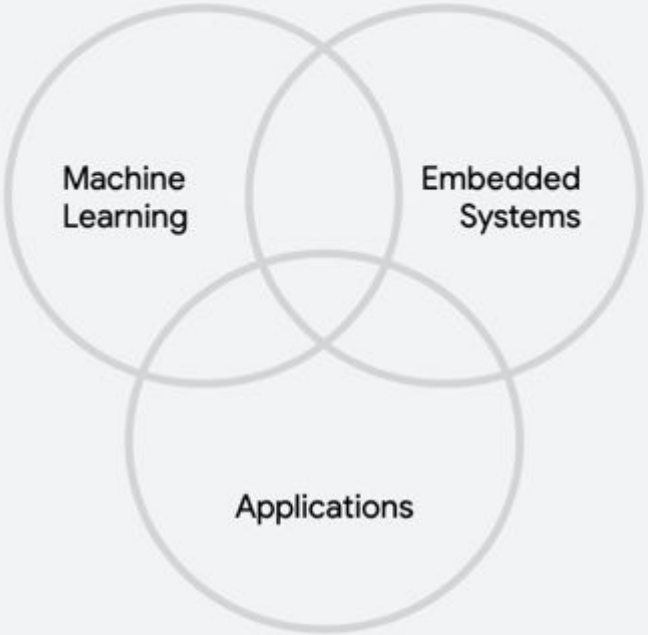
What areas we will learn?



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

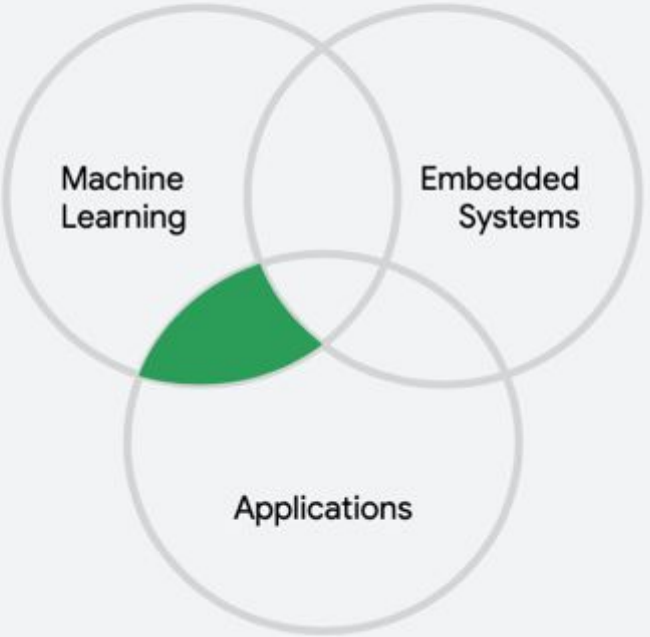
Interactions

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



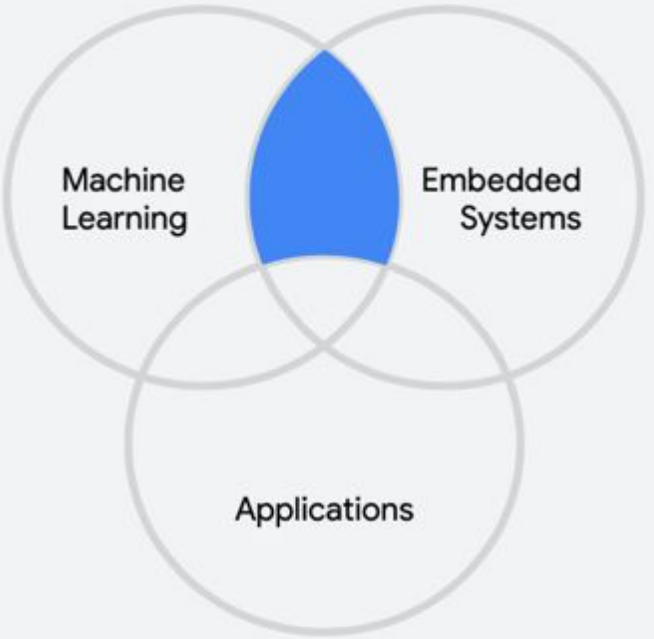
Interactions

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



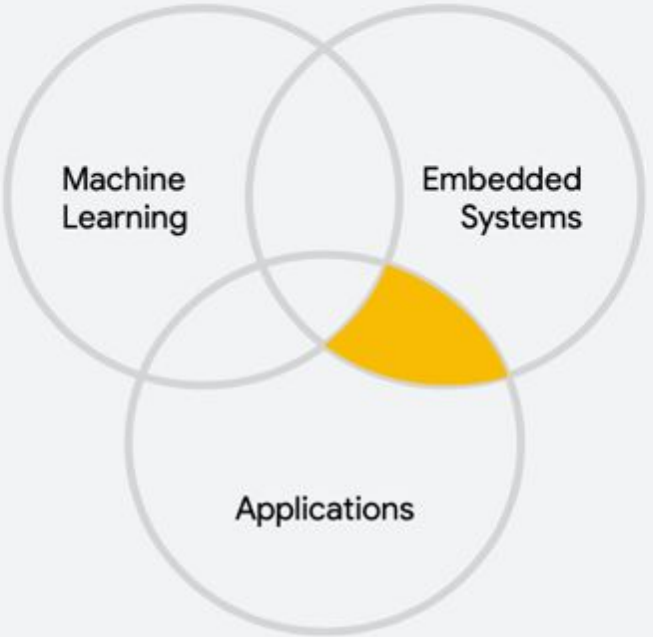
Interactions

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



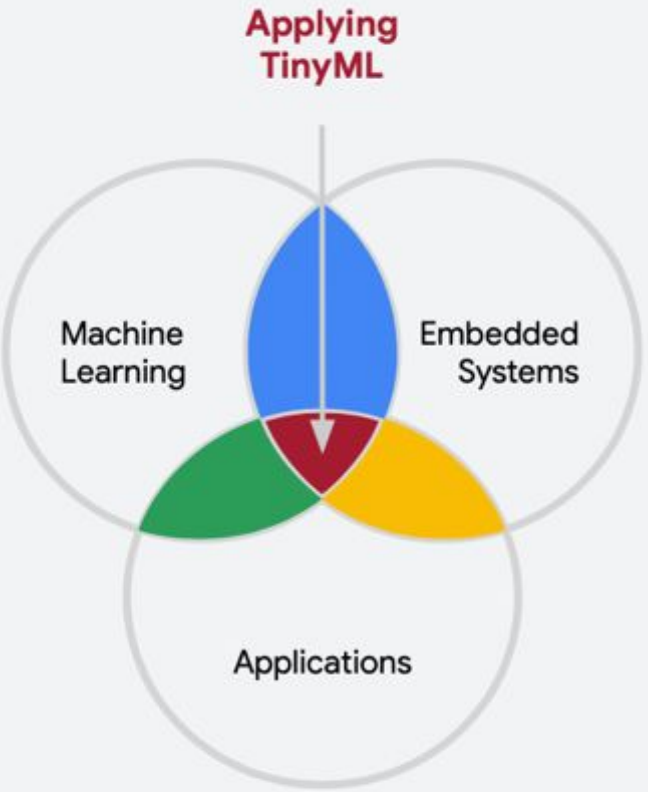
Interactions

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?

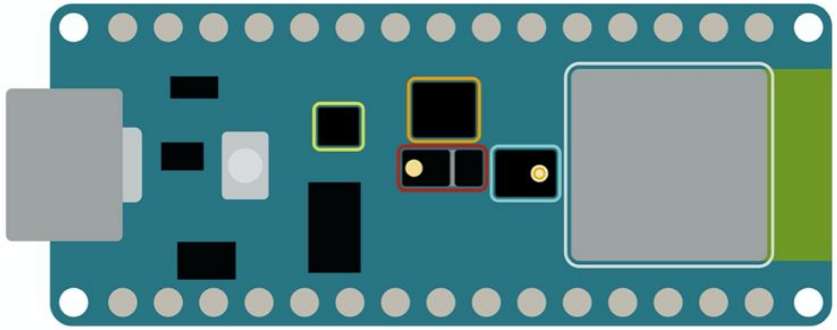
Hands-on Learning

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

TinyML Kit


ARDUINO
TINY MACHINE
LEARNING KIT
in partnership with:
edX
TensorFlow
OV7675

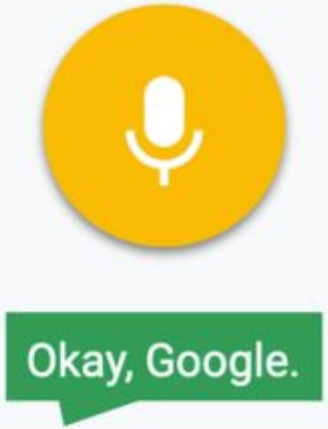
Nano 33 BLE SENSE & OV7675 Camera



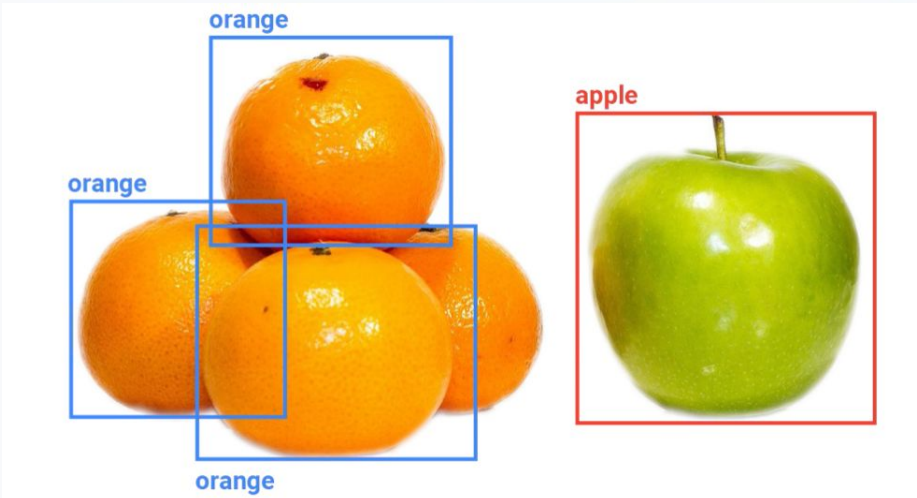
- Color, brightness, proximity and gesture sensor
- Digital microphone
- Motion, vibration and orientation sensor
- Temperature, humidity and pressure sensor
- Arm Cortex-M4 microcontroller and BLE module

Hands-on Activities

Speech



Vision

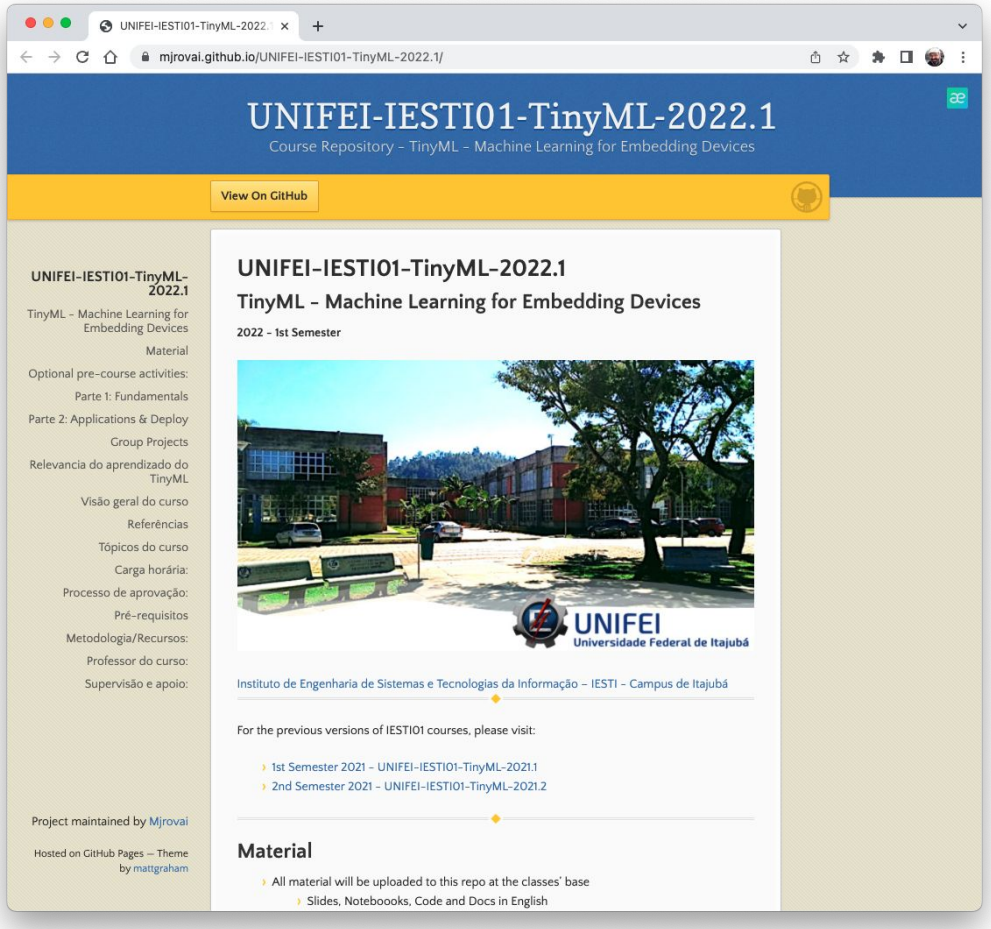


IMU

How is the course structured?

Course Structure

- Weekly live lectures (15 weeks)
  - Slides
  - Hands-on coding (by teacher & students)
- Weekly Additional Readings
- Guest Lecturer (*)
- Assignments
  - Quizzes (Weekly)
  - Notebooks with codes (5)
  - 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

https://mjrovai.github.io/UNIFEI-IESTI01-TinyML-2022.1/

Class planning and approval process

- Minimal suggested Workload (4 hours per week):
  - 30 hours (Weekly online classes of 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

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

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