Seamless TinyML lifecycle management

In Software Engineering Project with University of Helsinki CS

16/1/2023

Origami@NEXUS: Hiroshi Doyu, Roberto Morabito, Michihito Mizutani

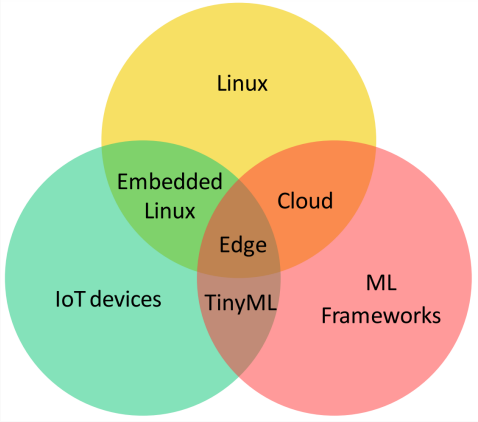
Who we are, Origami*

WHO WE ARE, Origami @ NEXUS

- NEXUS
  - 
  - Helsinki university incubation program
  - 6 month program: 9/2022-3/2023
  - Aiming at
    1. Start piloting
    2. Get funding
- Origami
  - Originally started by E/// Researcher x2
    - **Hiroshi Doyu** • You
      Embedded AI Researcher
      Espoo
      Experience: Ericsson, NVIDIA, and 5 more
      1 Eurentrepreneur, AI researcher, Kernel Hacker
    - **Roberto Morabito** • 1st
      Researcher and Teacher at University of Helsinki
      Helsinki Metropolitan Area
      Experience: University of Helsinki, Ericsson, and 5 more
      2 Senior Researcher @ Helsinki university
  - Joined by Ex Nokia colleague
    - **Michihito Mizutani** • 1st
      Design Lead at KONE
      Helsinki
      Experience: KONE, Aalto University, and 6 more
      3 Design Lead @ KONE
  - Mentored by
    - **Perttu Tuomaala** • 1st
      Vice President
      Porvoo
      Experience: Neste, Ensto, and 6 more
    - **Samuli Salmela** • 1st
      CIO, Digital leader - both in metaverse and physical world.
      Helsinki Metropolitan Area
      1K followers

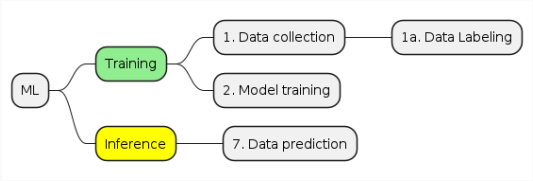
Project goal

*"The main goal of this software engineering project is to develop a solution that enables a seamless **TinyML lifecycle management**. In particular, the idea is to build a framework that **in an automated fashion** performs the different steps of the TinyML lifecycle management."*, from the original application

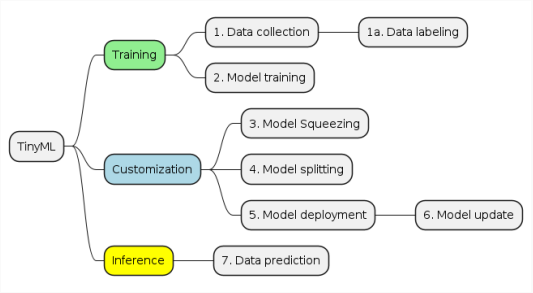


Lifecycle of: ML vs TinyML

(Cloud) ML



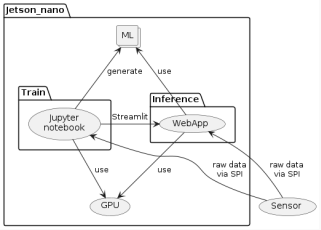
TinyML



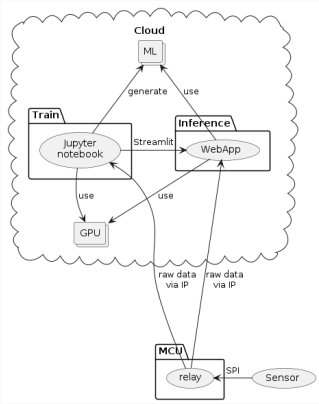
Arch: Edge ML vs Cloud ML vs TinyML

Edge ML

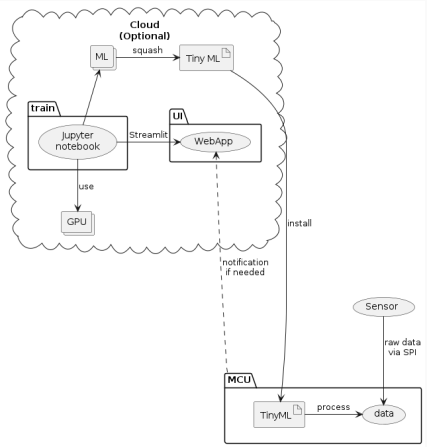
(Local ML)



Cloud ML



TinyML



TensorFlow Lite for Microcontrollers*

ML model Examples

- hello_world
- magic_wand
- memory_footprint
- micro_speech
- mnist_lstm
- network_tester
- person_detection

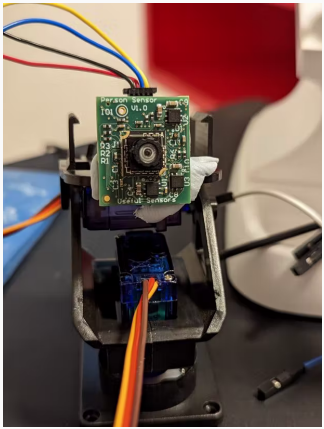
Supported platforms

TensorFlow Lite for Microcontrollers is written in C++ 11 and requires a 32-bit platform. It with many processors based on the Arm Cortex-M Series architecture, and has been porte including ESP32. The framework is available as an Arduino library. It can also generate pro environments such as Mbed. It is open source and can be included in any C++ 11 project.

The following development boards are supported:

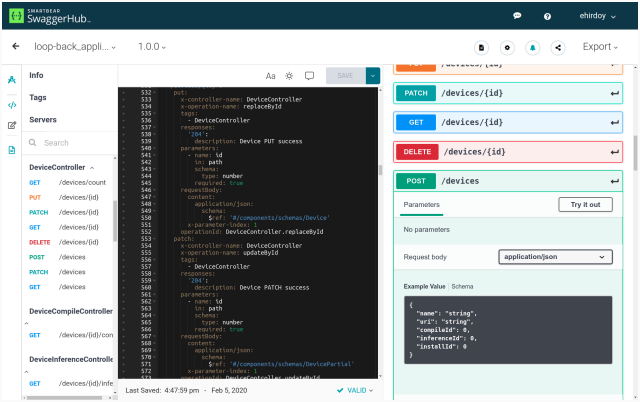
- Arduino Nano 33 BLE Sense
- SparkFun Edge
- STM32F746 Discovery kit
- Adafruit EdgeBadge
- Adafruit TensorFlow Lite for Microcontrollers Kit
- Adafruit Circuit Playground Bluefruit
- Espressif ESP32-DevKitC
- Espressif ESP-EYE
- Wio Terminal: ATSAMD51
- Himax WE-I Plus EVB Endpoint AI Development Board
- Synopsys DesignWare ARC EM Software Development Platform
- Sony Spresense

Face-Following Pan/Tilt Stand*

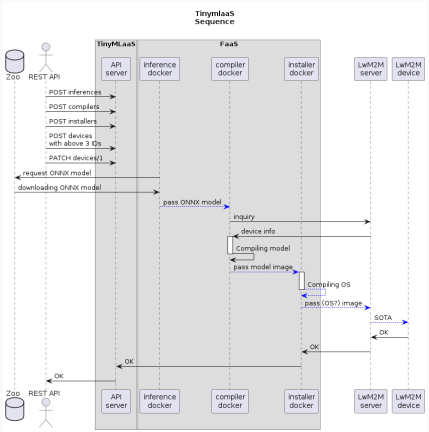


Automate lifecycle with TinyML as-a-Service API

OpenAPI spec for TinyMLaaS (Old)

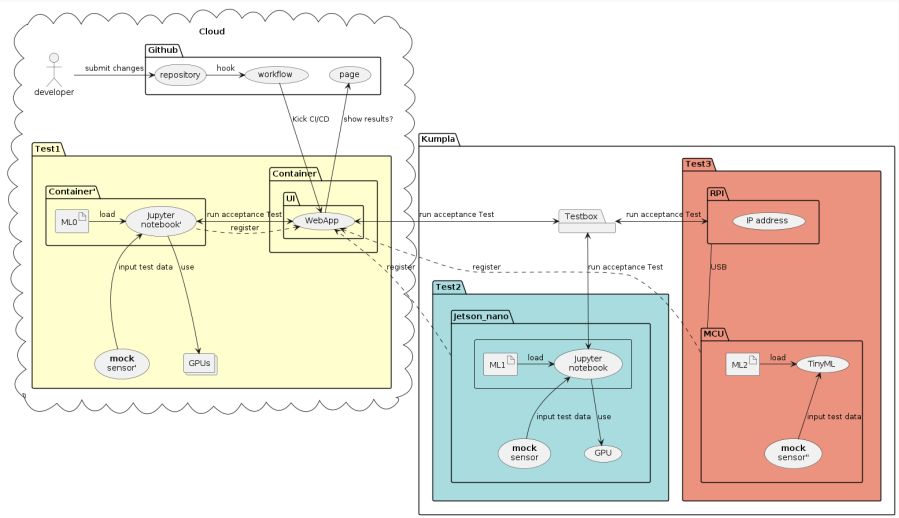


Function as-a-Service (FaaS)



TinyMLaaS orchestrates TinyML on ***any IoT system***.

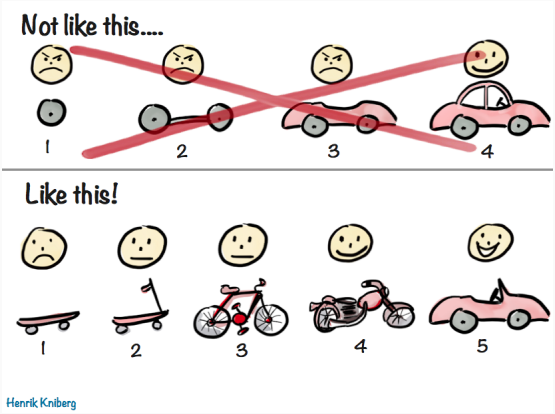
CI / CD / ATDD



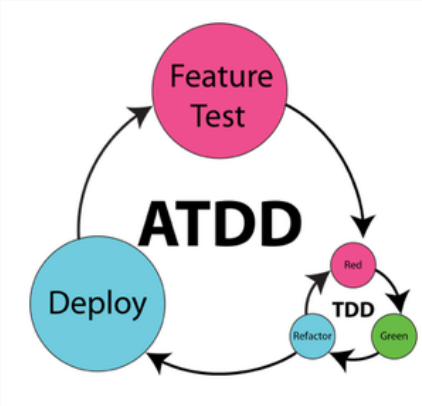
The simplest **Test1**: *TFLite micro Hello World* in x86 container w/o HW.

MVP iteration

Always runnable MVP at Day 1



Acceptance Test Driven Development



User story

As a [persona], I [want to], [so that]

- As a Data Scientist at training,
  - I want to collect data to train
    - I want to label data to train
  - I want to train models to use devices
  - I want to store models to assign
- As a on-site IT operator,
  - I want to register:
    - IoT devices to observe
    - models to update
    - toolchain to compile
  - I want control panel:
    - to assign models
    - to build ML pipelines
- As a CFO,
  - I want to compare Cloud vs TinyML for cost
  - I want to pipeline Cloud & TinyML for flexibility
- As a CEO,
  - I want dashboard to observe devices
- As a CTO,
  - I want automated dry-run of a whole lifecycle to reject support requests

.... e.t.c

Contact information

Origami

https://Origami-TinyML.github.io/blog/about.html