tinyML for Good

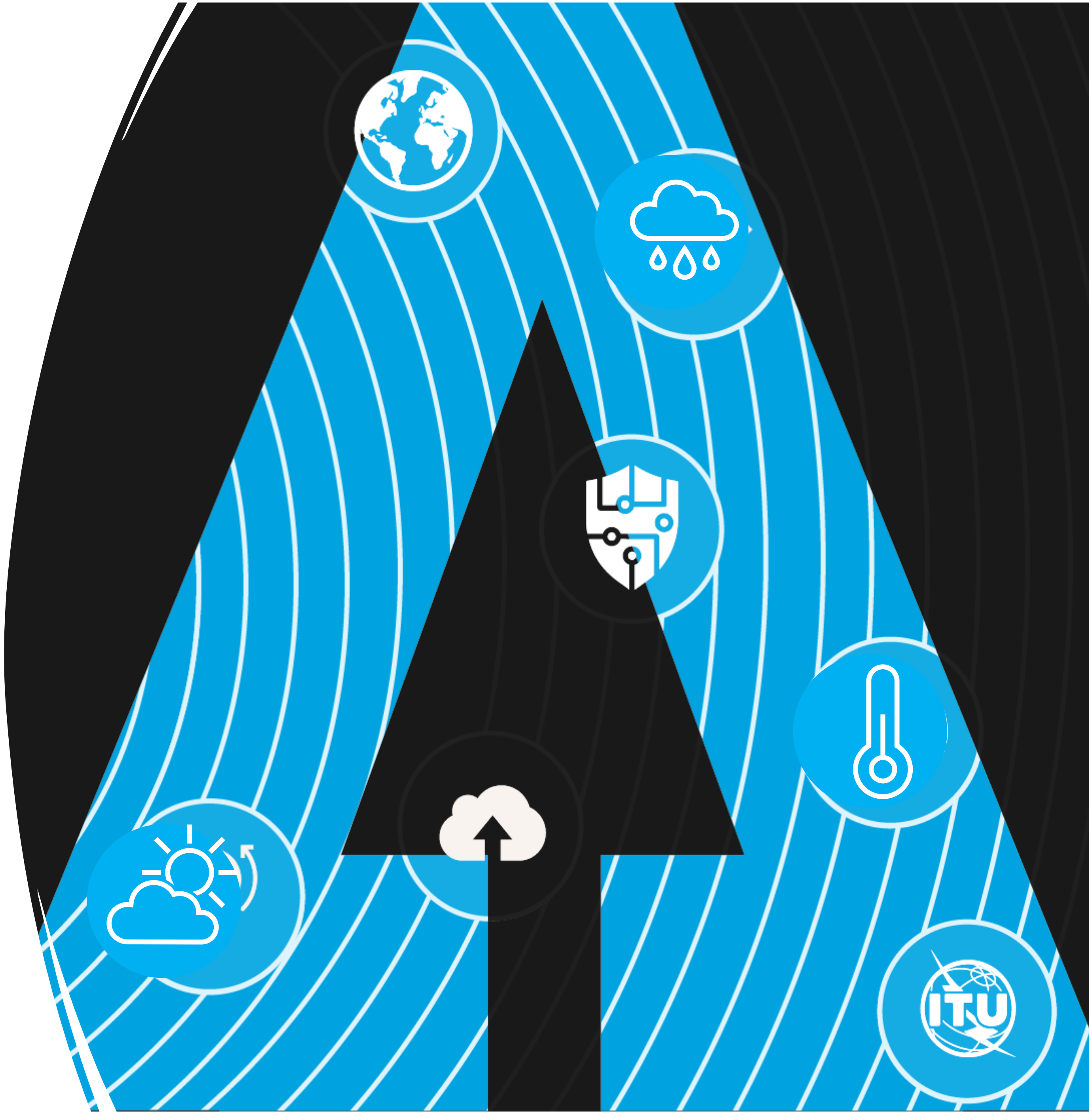
The role of tinyML in advancing UN SDGs

ICTP-UNU tinyML workshop

Thomas Basikolo, PhD

International Telecommunication Union (ITU)

26 April 2024



Agenda

1. Background

- ITU (UN agency for digital)
- tinyML for Good

2. tinyML for Good & use cases

- Next-Gen Smart weather station
- Pedestrian detection
- Crop disease and wildlife

3. Challenges or Issues [tinyML & AI]

- Ethics
- Regulation and governance
- Sustainability
- e-waste

International Telecommunication Union (ITU)

The UN specialized agency for ICT

ITU-R

Allocation of radiofrequency spectrum and satellite orbits

ITU-D

Bridging the digital divide

ITU-T

Establishing international standards

Standards development in ITU

- Developing **technical standards** for the interconnection and interoperability of digital technologies
- Bridging the **standardization gap** between developed and developing countries
- Fostering cooperation among national, regional and international **standards bodies**

Study Groups

Focus Groups

Workshops and symposia

Standards development in ITU: Technical foundations

Transport, access and home networks

Multimedia

Service quality

Numbering & emergency comms

Artificial intelligence

Cybersecurity

Internet of Things

Environmental efficiency

Quantum information tech

Accessibility

ITU Events

AI for Good

Accelerating the United Nations Sustainable Development Goals

ALL YEAR
ALWAYS **ON**LINE

aiforgood.itu.int

38 UN PARTNERS

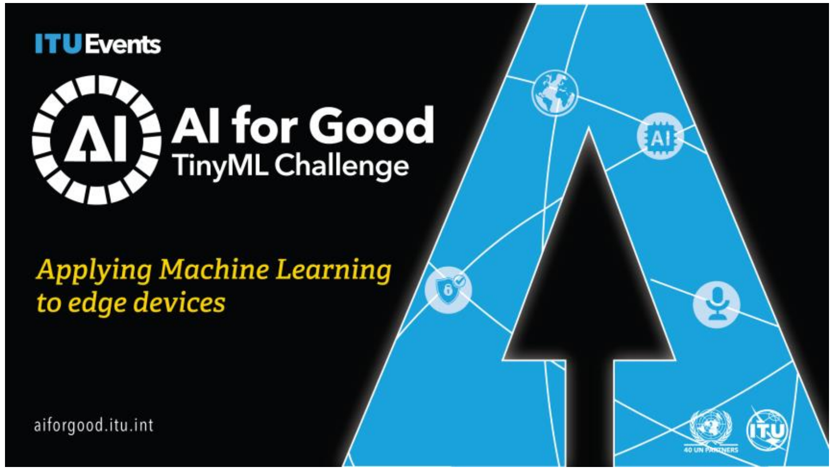
AI for Good | The goal

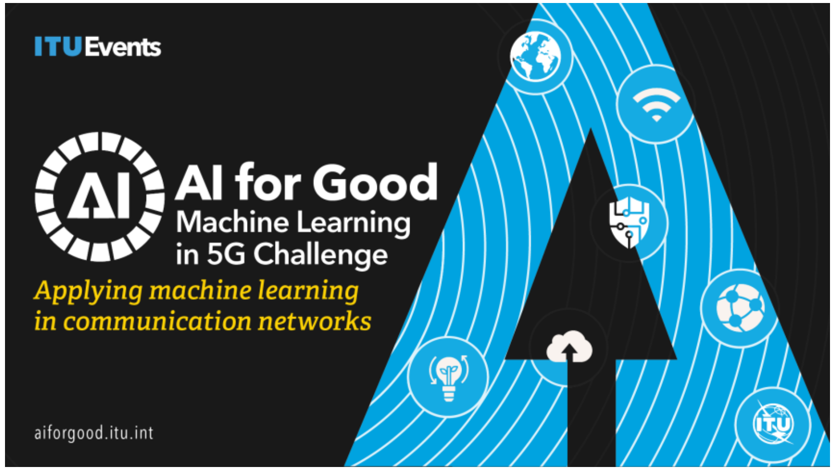
Identify
practical applications of AI

Scale
solutions for global impact

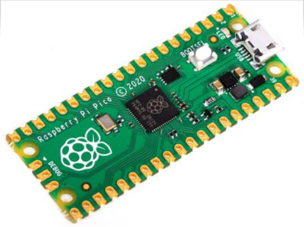
Accelerate
progress towards the UN SDGs

ITU AI/ML Challenges



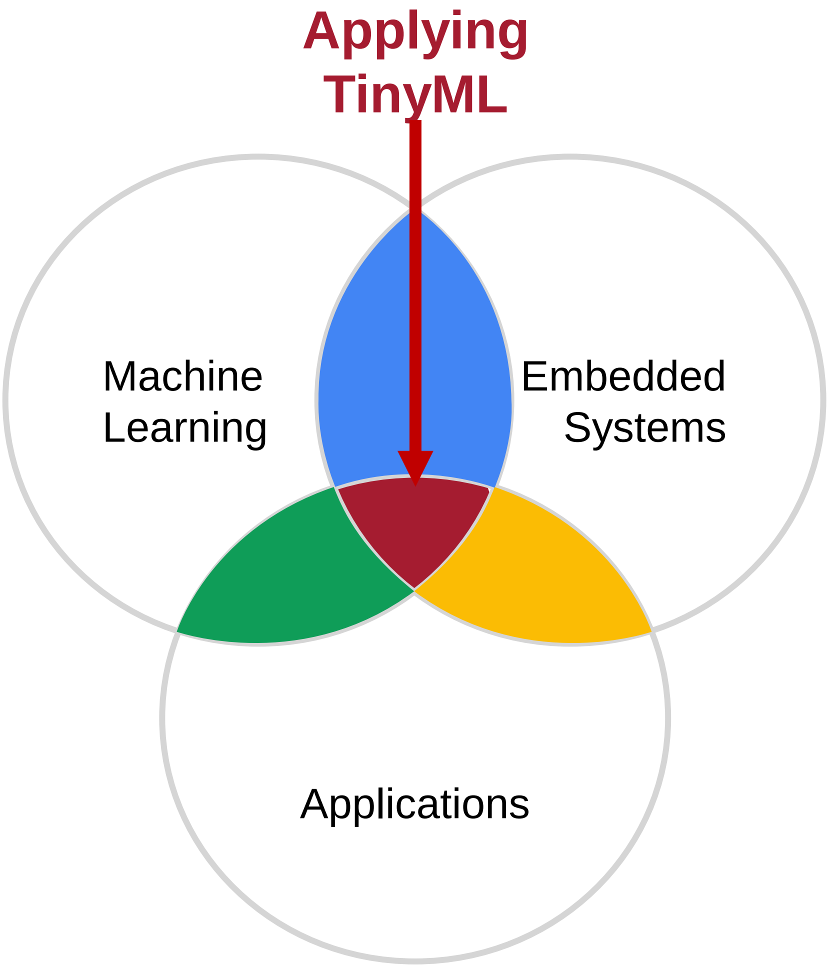


tinyML for Good



tinyML

- ML enabling new and interesting TinyML applications
- Challenges with enabling machine learning on tiny, resource-constrained embedded devices
- New use cases can we possibly enable on embedded systems



Benefits of tinyML

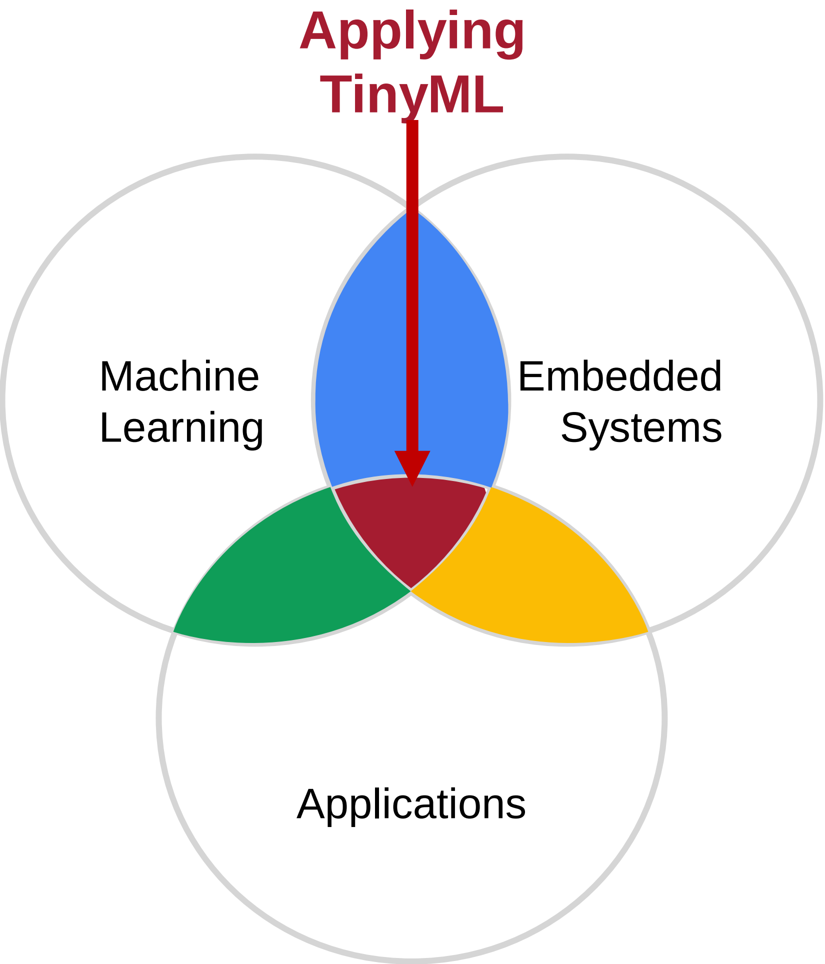
- Energy consumption - executing ML on low-power IoT devices significantly reduces energy consumption compared to running these algorithms on cloud servers.

- Privacy and security - collecting and analyzing data on IoT devices without exchange with external servers ensures privacy & data security.

- Real-time decision-making

- Autonomy and efficiency - enables IoT devices to make autonomous decisions based on real-time data analysis

Mapping tinyML to SDGs

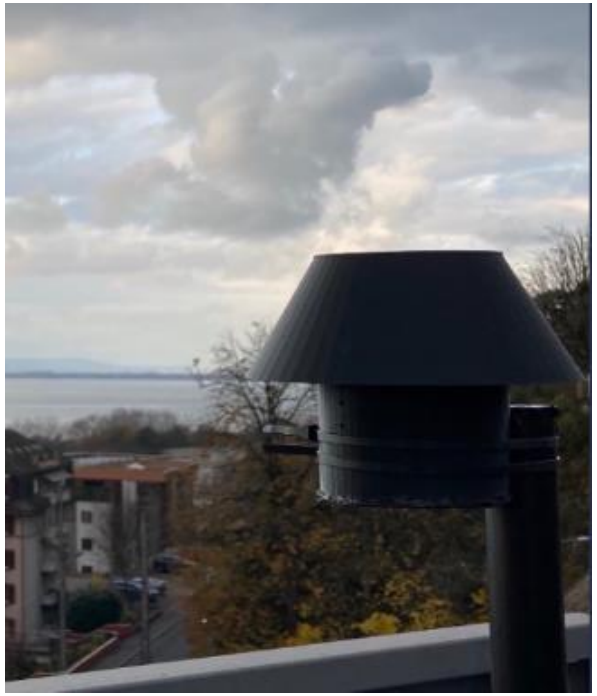


tinyML for (social) Good

tinyML for Good

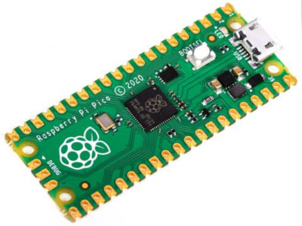
- Climate Change
- Health
- Conservation
- Energy
- Food Security

Approx. $3'000



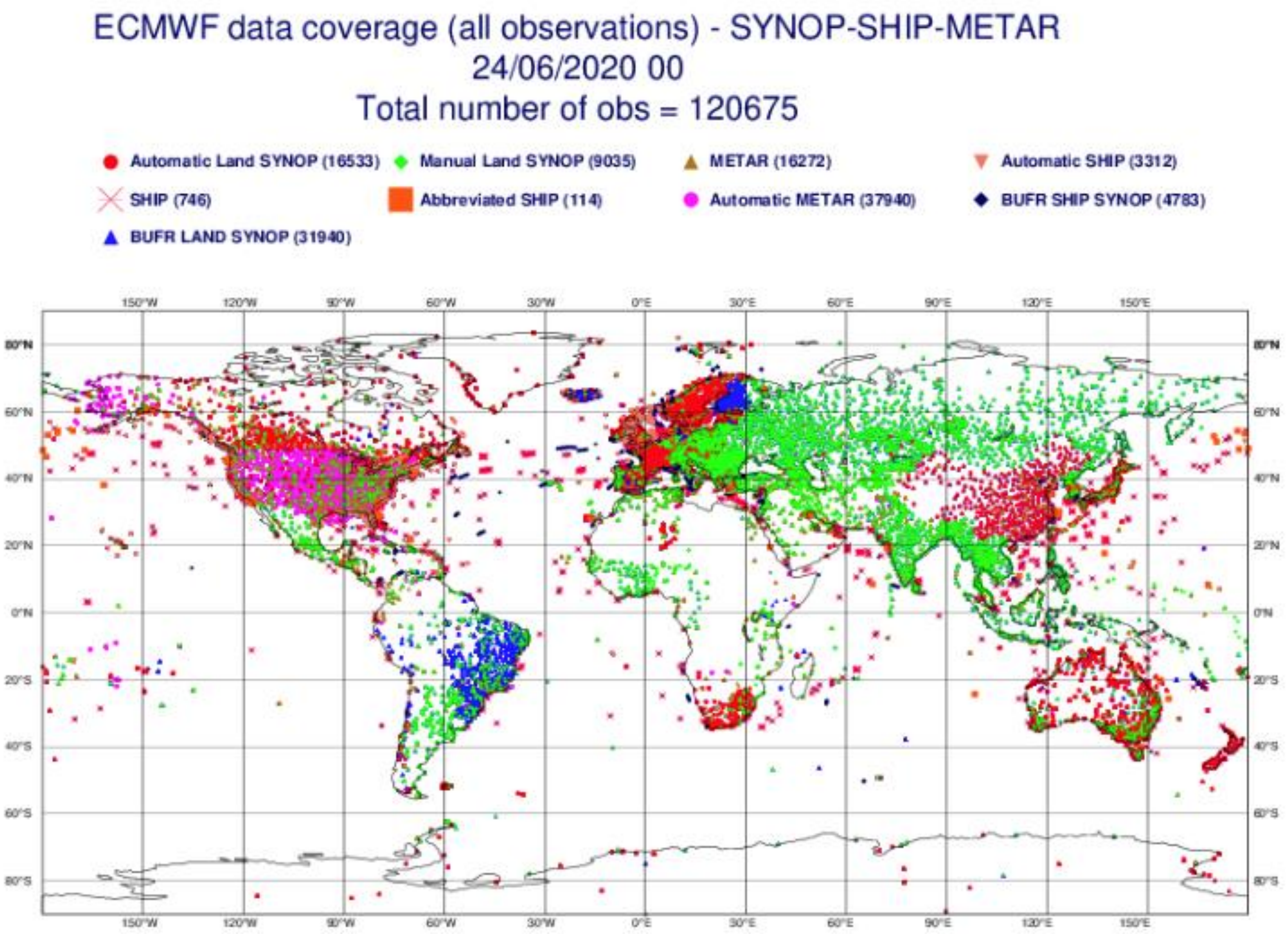
$75

tinyML use cases



tinyML Challenge: Smart weather station

Developing Countries is the area of the globe where land-based, in situ monitoring of weather and climate is at its scarcest, but at the same time has arguably the most potential to benefit society



credit: Marco Zennaro [ICTP]

tinyML Challenge: Smart weather station

CLIMATE CHANGE

54% of Africa's surface weather stations can't capture data properly

With extreme weather events becoming the new normal, forecasting has become more important than ever. It's time African nations invest in observation networks

NEXT NEWS >

By Maina Waruru
Published: Tuesday 29 May 2018

AFRICA

A lack of weather data in Africa is thwarting critical climate research

By Rachel Chason and Rael Ombuor
September 24, 2021 at 5:00 a.m. EDT

credit: Marco Zennaro [ICTP]

tinyML Challenge: Smart weather station

The goal was to create a **low-cost**, **low-power**, **reliable**, **accurate**, **easy to install** and **maintain weather station**, with **no mechanical moving parts** for measuring all weather conditions with a focus on rain and wind measurement, **based on ultra-low power ML at the edge**.

Wind and rainfall measurements – existing solutions

Anemometer

Tipping bucket

+ Simple
+ Easy to implement
- Moving parts
- High cost
- Prone to failure

What we offered to participants

- A community/network
- Use case with real effect
- Impact and contributing to SDGs
- Possibility to create a full stack solution
- Mentorship

Webinars

- ITU & ICTP
- Edge Impulse
- Syntiant
- University Corporation for Atmospheric Research (UCAR)
- Renesas

2022 Challenge Prizes and awards

Cash Prize

$3’000

$2’000

$1’000

Special Prizes/awards

- Best use of Edge Impulse prize $3,000
- Best use of Syntiant TinyML Board $500
- Additional 1st place prize: SenseCAP S1000 10-in-1 Compact Weather Sensor with CO2 Measurement
- Best vision-based weather station implementation - $500 in Sony merchandise or gift card

WINNER CERTIFICATE

tinyML Challenge

Gold Champion (1st prize)

is presented to

Team tinyML

for their innovative solution to the problem statement "Smart weather station" in the ITU AI/ML in 5G Challenge 2022

14 December 2022

Houlin Zhao
Secretary General, ITU

Chaesub Lee
Director, ITU Standardization Bureau

EDGE IMPULSE

SYNTIANT

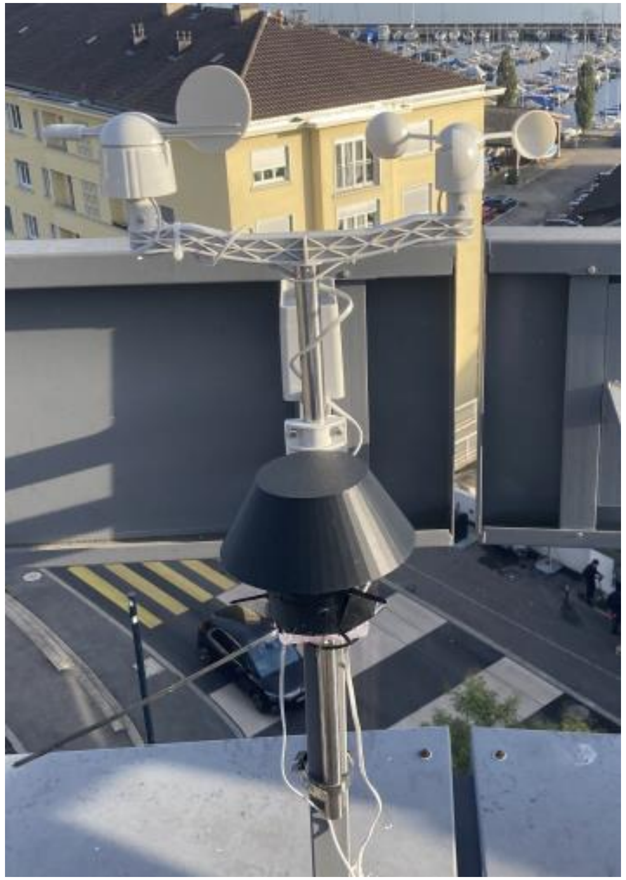
seeed studio
The IoT Hardware Enabler

Sony Semiconductor Solutions Corporation

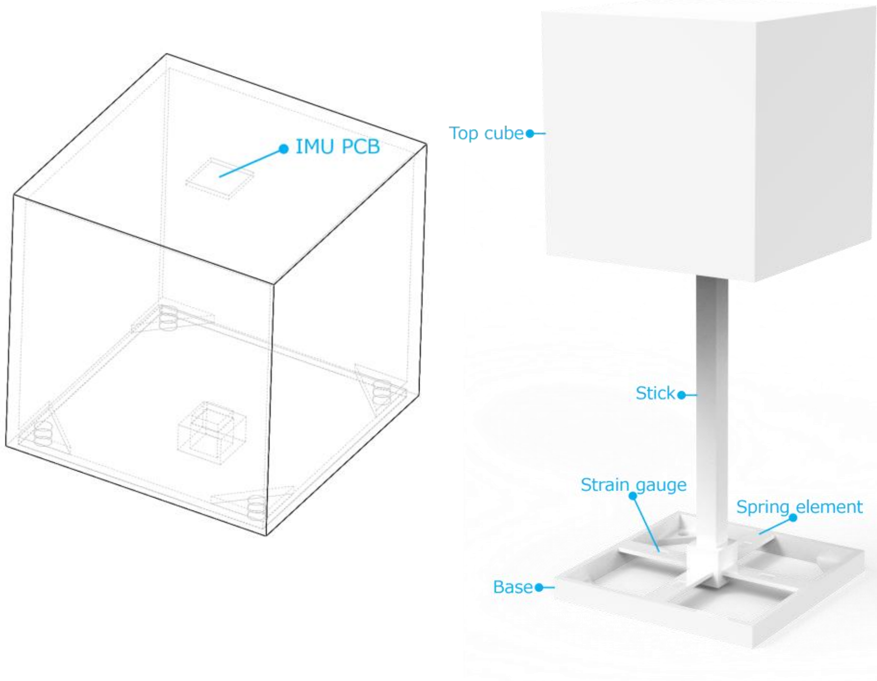
tinyML Challenge 2022: solutions

Winning solutions

CSEM: Aurora weather station

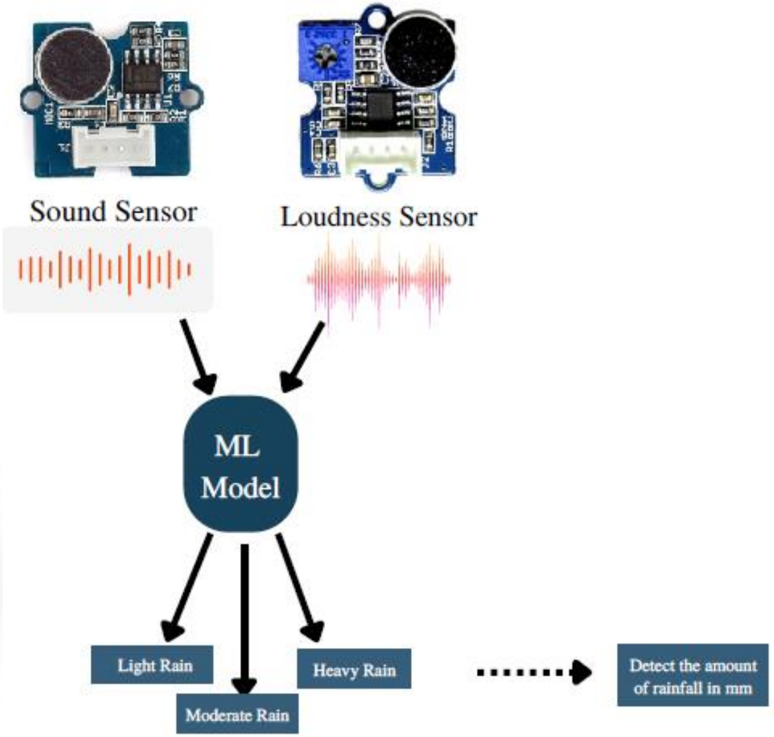


BrainCircuit



Room to grow and make impact

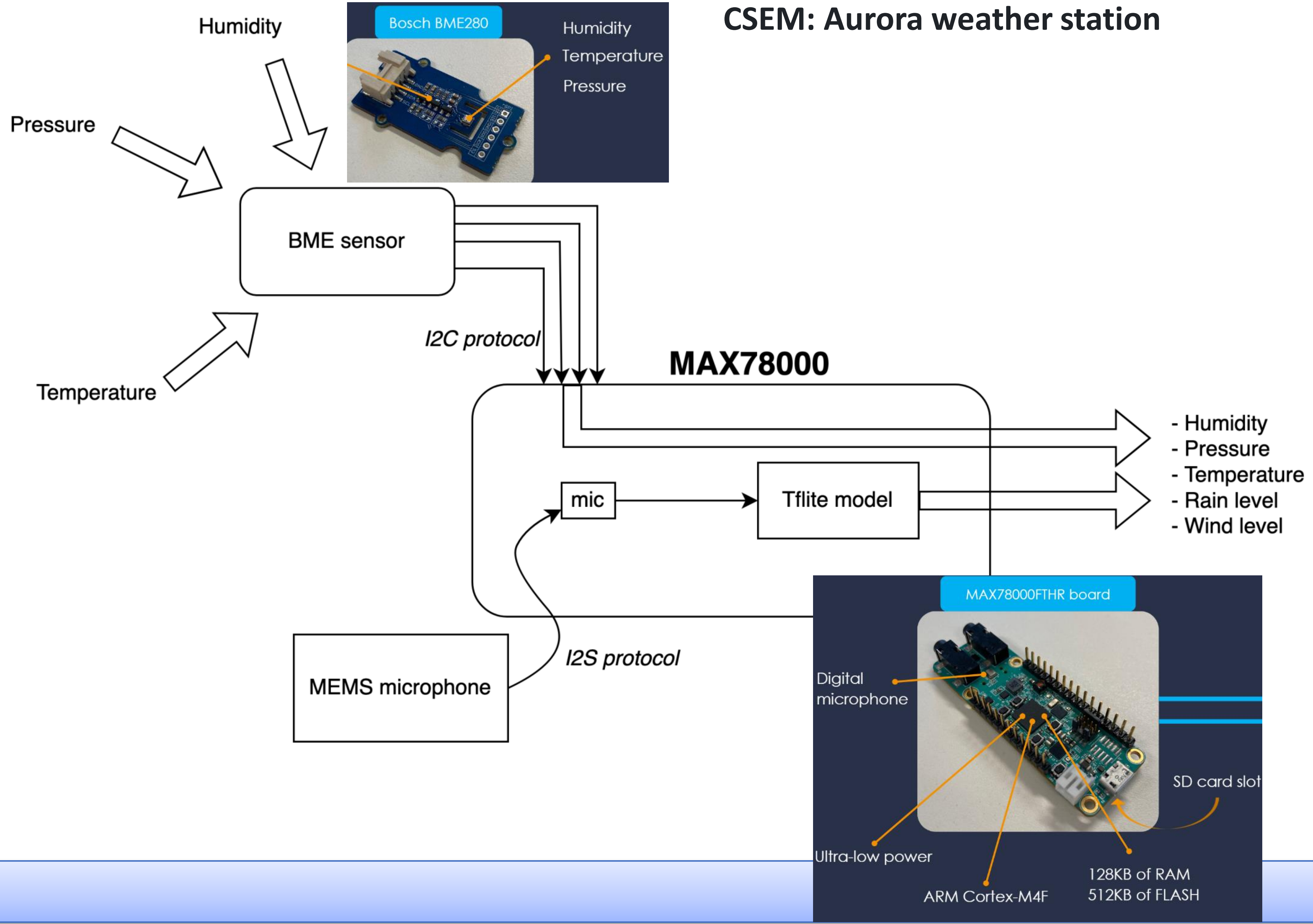
ICFOSS

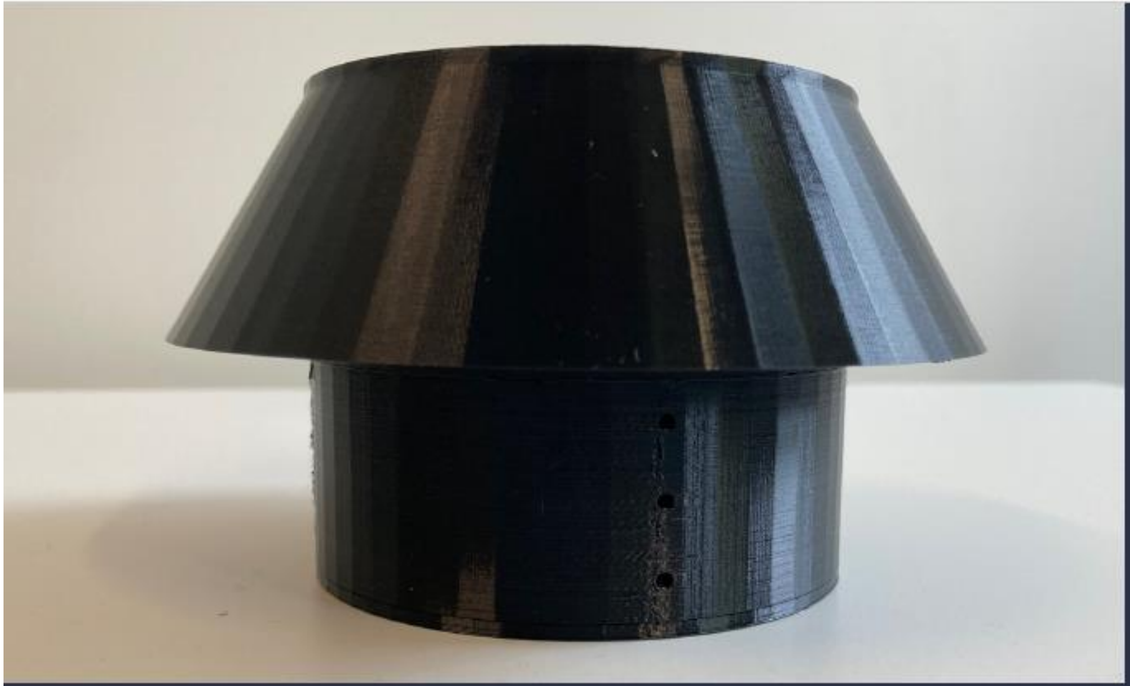


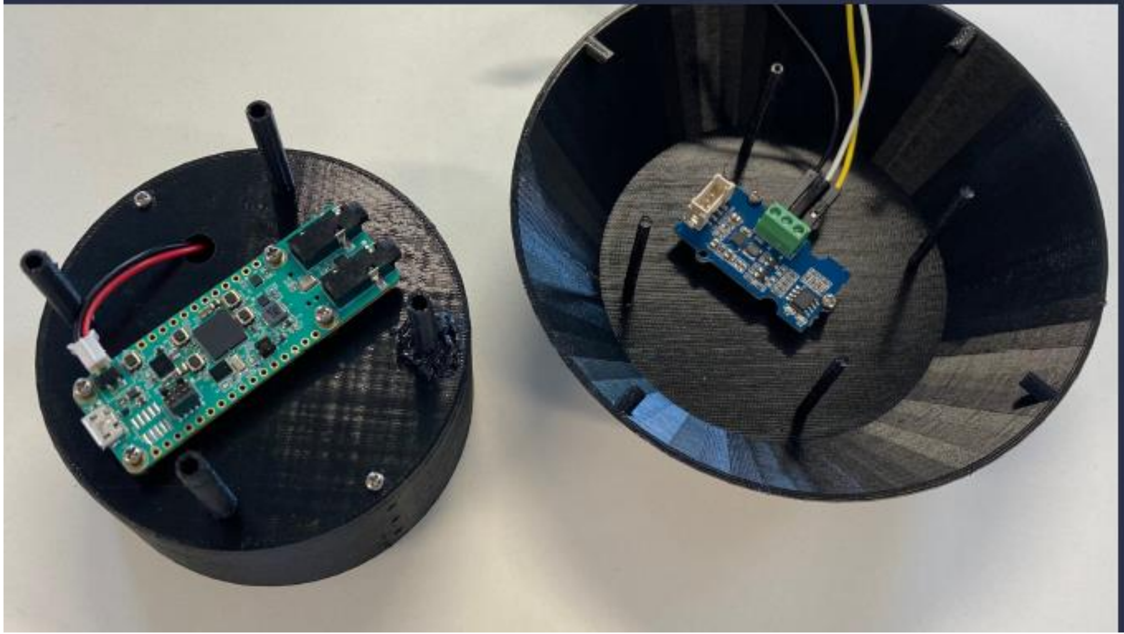
** Most sustainable product

tinyML Challenge 2022: solutions

CSEM: Aurora weather station





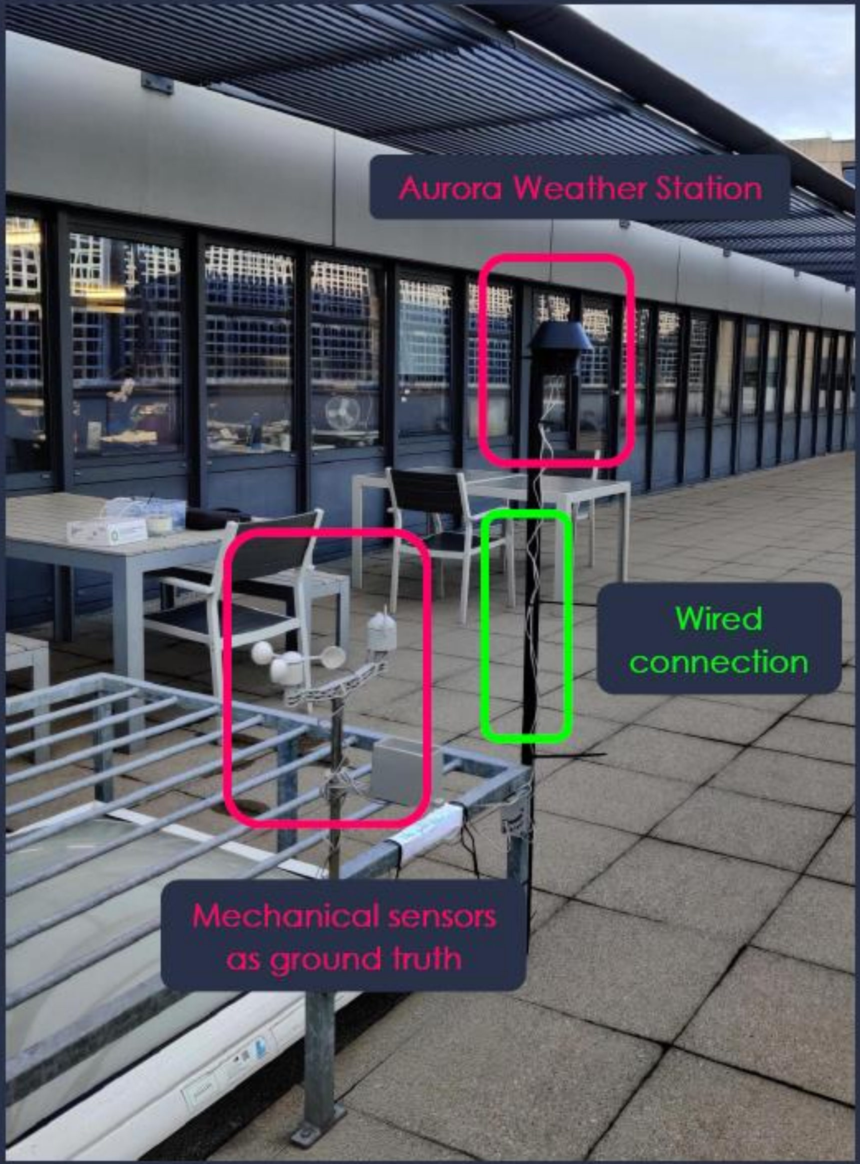


Aurora summary

- ✓ Detects wind and rain
  - 84.9% of accuary for rain, 82.1% for wind
- ✓ Maintenance-free
- ✓ Works in real-world conditions
- ✓ ML fully embedded (<150KB Flash)
- ✓ Easy to install

Update on Aurora

- Implemented Quantization-aware training (QAT) to minimize accuracy drop due to quantization
- Integration of mechanical sensors in the test setup (2 separate entities before)
  → Better synchronization ground truth and embedded sensors (microphone + env sensor)
- Sensor fusion of microphone with environmental sensor data
  →No major improvement



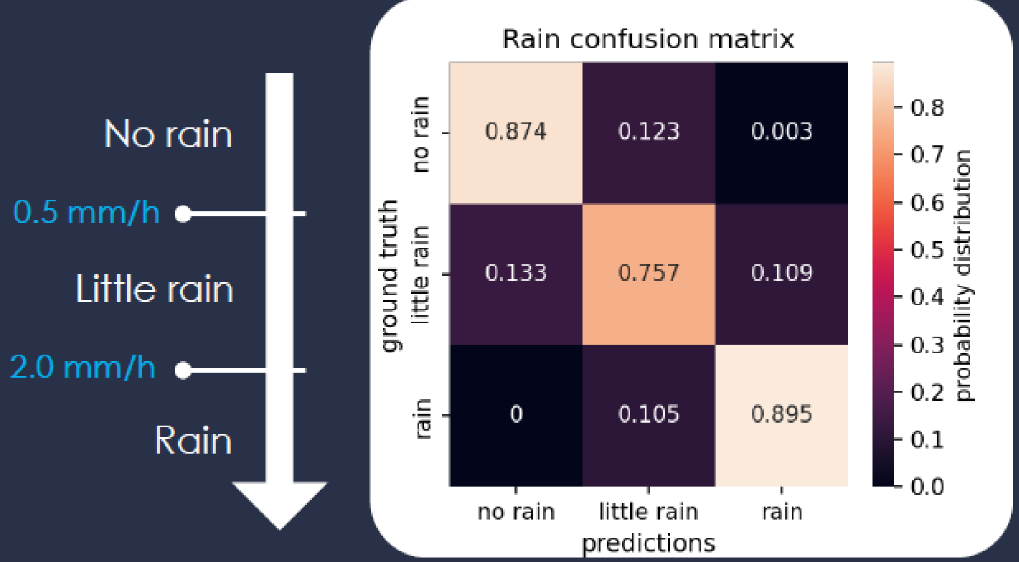
csem

CONFIDENTIAL

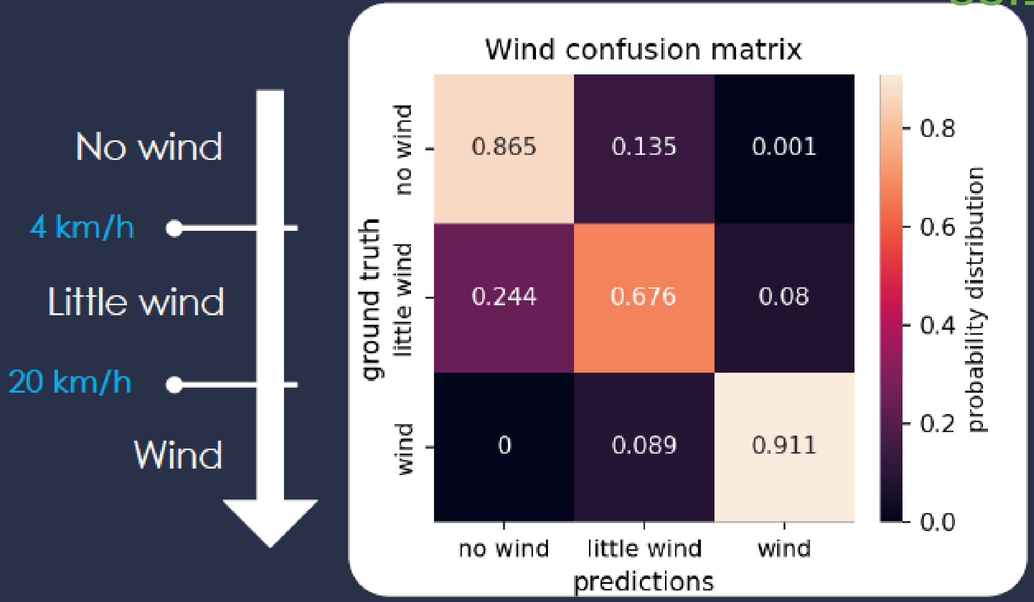
* Balanced accuracy of FP model on real-world test set

New results after improved labelling

- Rain
  - 3 classes : 84.9%* ⇦ 74.4% (Dec. 2022)
  - 2 classes (no+little rain, rain) : 94.0%* 85.7%



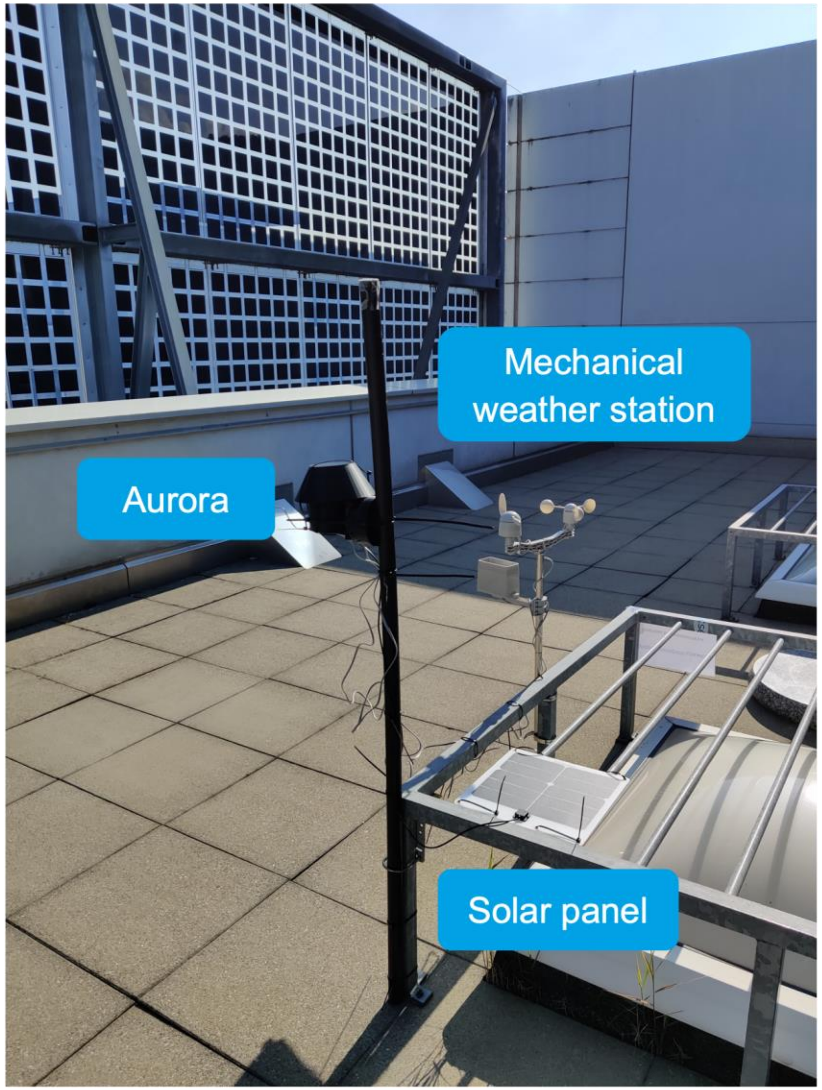
- Wind
  - 3 classes : 82.1%* ⇦ 70.5%
  - 2 classes (no+little wind, wind): 92.5%* 83.1%



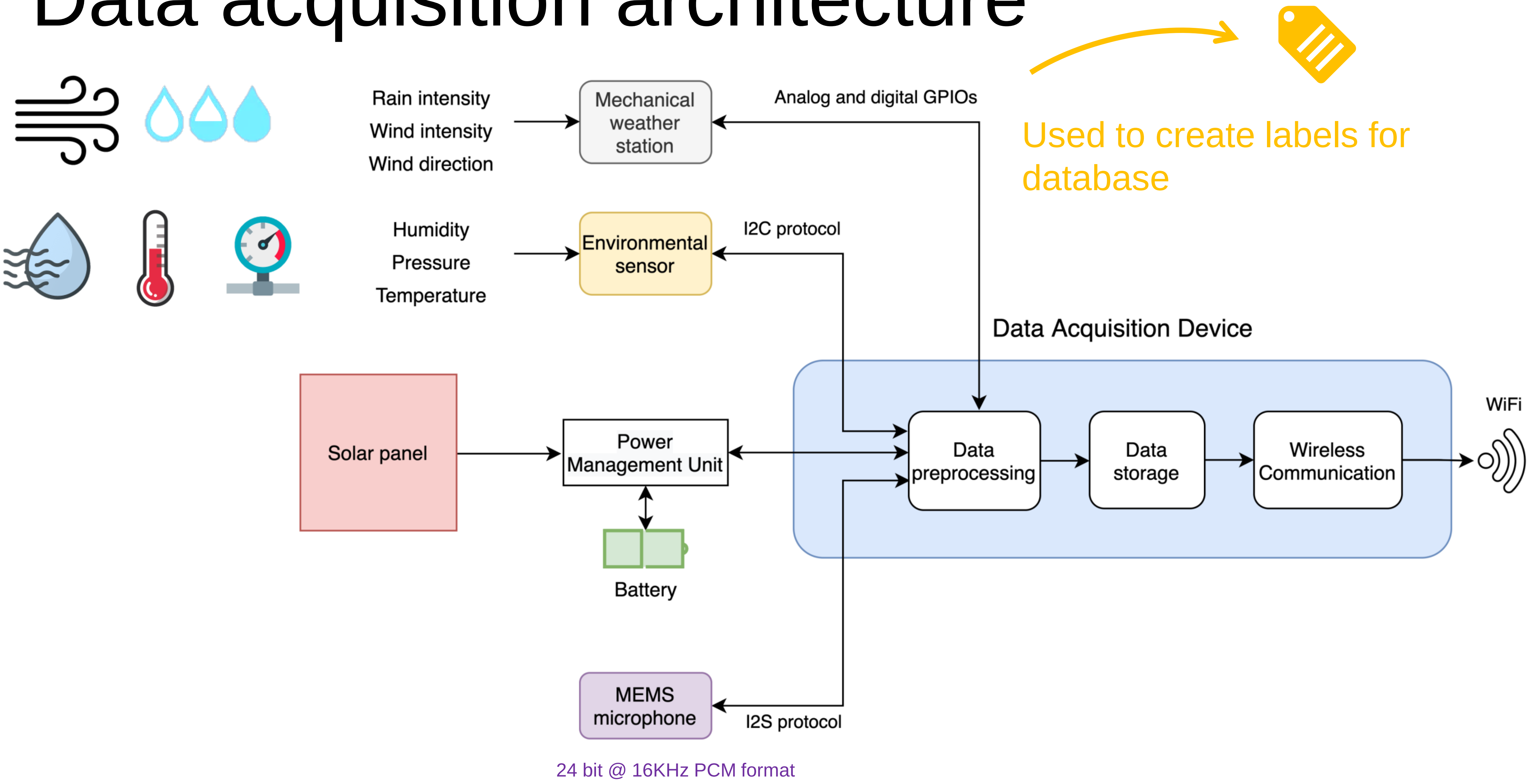
- Quantized model: loss of 3-5% after QAT

Next-Gen tinyML Smart Weather Station

- Synchronization between Aurora and reference station (mechanical weather station)
- Improved readout of sensors of reference station
- Solar-powered autonomous operation
- Wireless communication by WiFi, avoiding need to replace SD cards



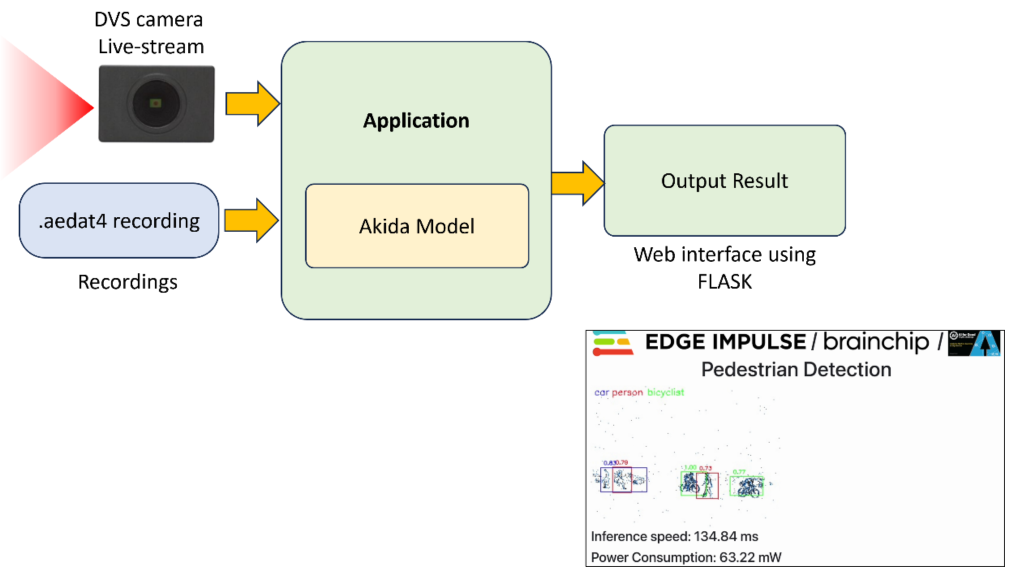
Data acquisition architecture



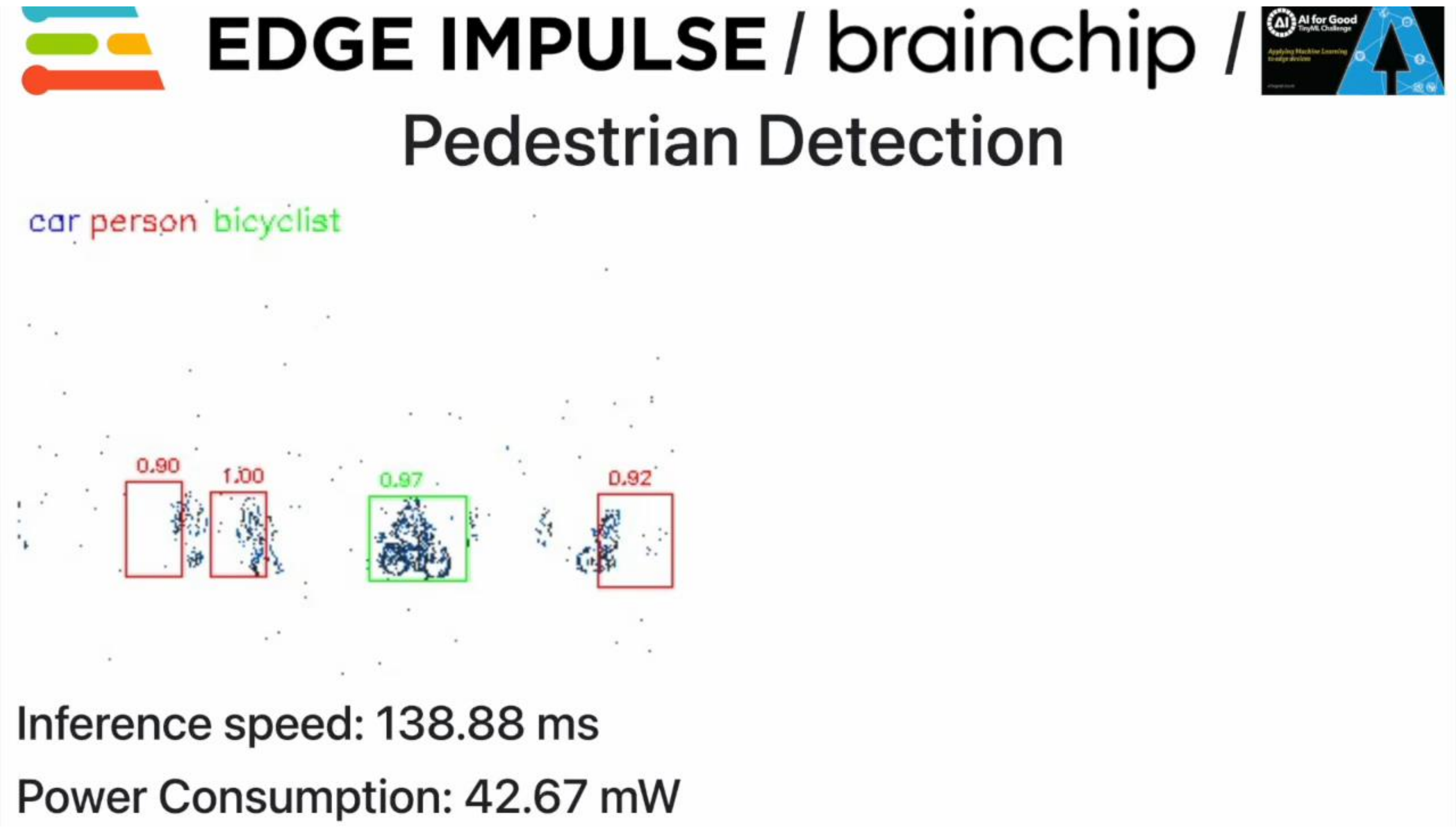
Pedestrian Detection

- In the US, traffic fatalities involving people walking have increased fast since 2009
- Devices to improve pedestrian detection usable in both day & night, placed intersections
- Challenges: people walking are small, and sometimes hard to see and differentiate from bicyclists, particularly at night
- Vision Zero: to reduce traffic fatalities and severe injuries

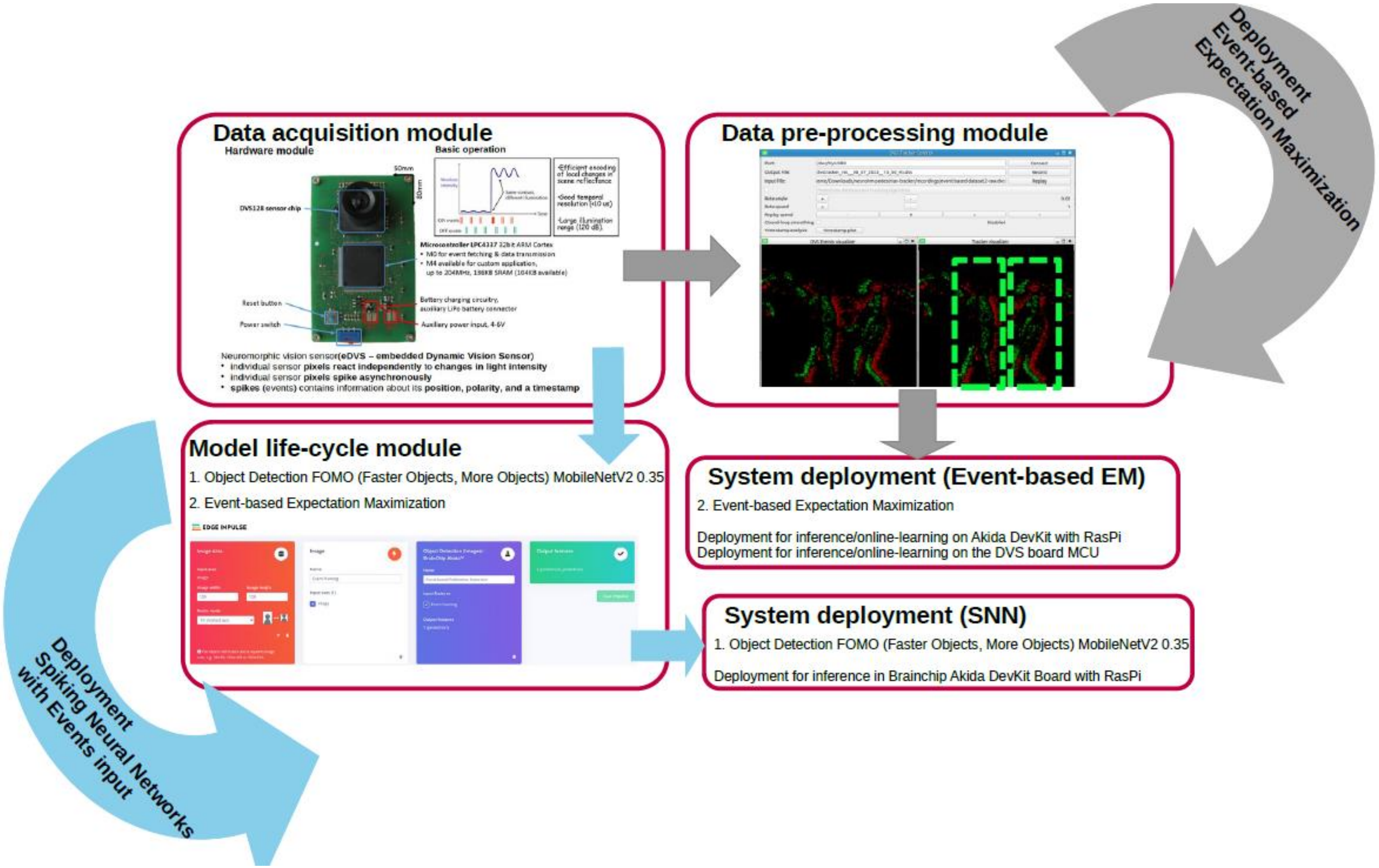
Pedestrian Detection



Running inference on Raspberry PI 4 with AKD100 mini PCIe developme

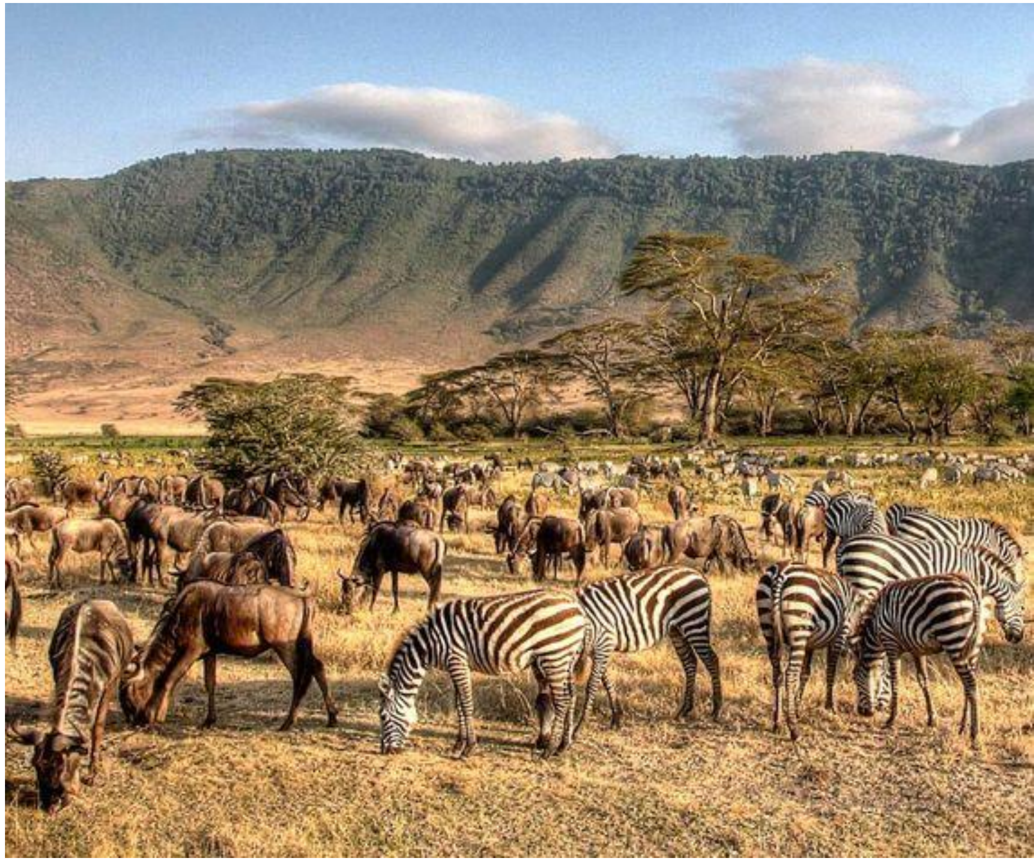


Pedestrian Detection



Wildlife monitoring

- Animal diversity is declining at an alarming rate (48%)
- Wildlife is necessary for prevention of drought, new deserts, fires as well as flood
- Management and conservation of animals carried out by human field workers
    - time-consuming
    - labor-intensive
    - expensive
- Camera traps are among the most used



Serengeti National Park, Tanzania

Wildlife monitoring

To develop a wildlife monitoring tool using TinyML classifying wild animals in limited resources areas

Live classification on ARDUINO_NANO33BLE

This shows a live preview of the camera on your device. Classification is running on the device, not on this computer.

deer (1.00)

Time per inference: 646 ms.

	BUFFALO	DEER	ELEPHANT	GIRAFFE	HORSE	LION	MONKEY	RHINO	TIGER	ZEBRA
65	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
64	0.00	**0.85**	0.00	0.10	0.00	0.00	0.00	0.00	0.04	0.02
63	0.00	**0.80**	0.02	0.08	0.00	0.00	0.04	0.00	0.00	0.05

From Edge Impulse Studio

Using MobileNetV2

Model Model version: Quantized (int8)

Last training performance (validation set)

ACCURACY 96.3%

LOSS 0.14

Confusion matrix (validation set)

	BUFFAL	DEER	ELEPHA	GIRAFFE	HORSE	LION	MONKEY	RHINO	TIGER	ZEBRA
BUFFALO	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%
DEER	0%	98.1%	0%	0%	0%	0%	1.9%	0%	0%	0%
ELEPHAN	0%	0%	97.6%	0%	2.4%	0%	0%	0%	0%	0%
GIRAFFE	0%	2.4%	0%	97.6%	0%	0%	0%	0%	0%	0%
HORSE	0%	2.1%	0%	0%	93.6%	2.1%	0%	2.1%	0%	0%
LION	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%
MONKEY	0%	0%	3.2%	0%	0%	0%	96.8%	0%	0%	0%
RHINO	5.6%	0%	5.6%	0%	0%	0%	0%	88.9%	0%	0%
TIGER	0%	0%	0%	0%	0%	2.1%	0%	0%	97.9%	0%
ZEBRA	0%	0%	2.3%	0%	0%	2.3%	0%	0%	2.3%	93.0%
F1 SCORE	0.96	0.97	0.94	0.99	0.96	0.96	0.97	0.93	0.98	0.96

Enal

Quantized (int8) Selected

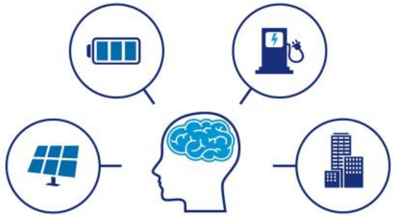
	IMAGE	TRANSFER LEARNING	TOTAL
LATENCY	11 ms.	3,758 ms.	**3,769 ms.**
RAM	4.0K	254.0K	**254.0K**
FLASH	-	307.1K	**-**
ACCURACY			**89.28%**

Unoptimized (float32) Select

	IMAGE	TRANSFER LEARNING	TOTAL
LATENCY	11 ms.	12,069 ms.	**12,080 ms.**
RAM	4.0K	816.3K	**816.3K**
FLASH	-	870.9K	**-**
ACCURACY			**92.27%**

Expected On-device performance metrics after deployment

Other application / use cases



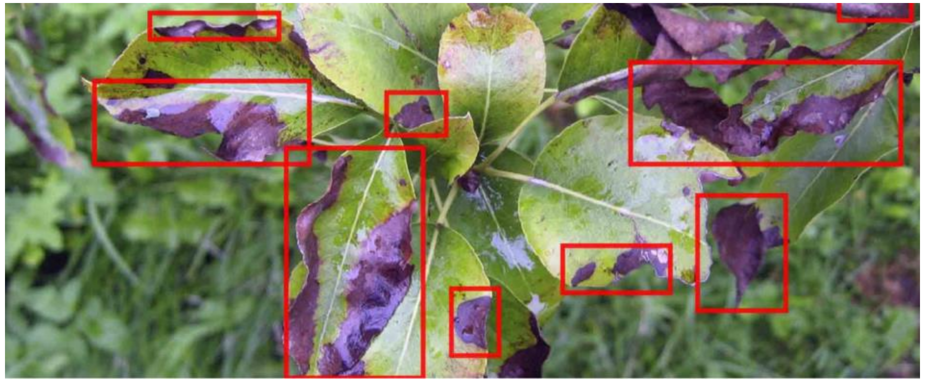
seeed studio The IoT Hardware Enabler ARDUINO EDGE IMPULSE csem ICTP

Sum all: A new perspective on tinyML

- Well-defined/standardized datasets that represent real-world scenarios
- Real-world deployment challenges that tinyML systems may face
- Open collaboration
- Consider ***for Good*** applications or scenarios

Next-Gen tinyML Smart Weather Station


Wildlife Monitoring


Plant Disease Detection

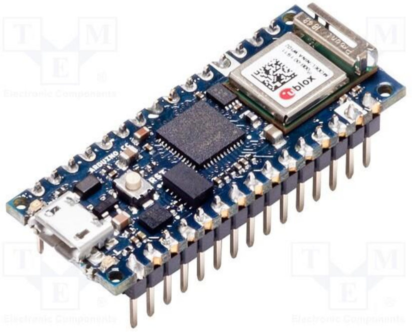
Challenges of tinyML

1. Memory & processing limitations

- IoT devices have limited memory and processing resources (hard to execute ML)
- Balancing model complexity with available resources

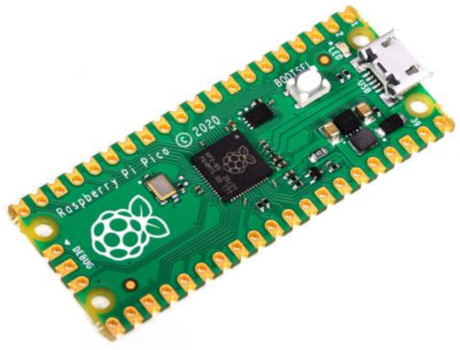
2. Training ML models

- Difficult to train models on resource-constrained devices
- Energy-efficient training techniques needed to optimize model performance without draining the device's battery



3. Variations in sensor quality (app. rely on sensor data)

- Sensor quality can vary significantly, affecting the reliability and accuracy of ML models
- Robustness to noisy or inconsistent sensor inputs is essential



4. Maintenance & updates

- Deployed TinyML models require regular maintenance

5. Interpretability

Challenges of tinyML & AI: Ethics

Key ethical considerations

- **Privacy**: Ensuring that data collected by tinyML devices is used responsibly and does not infringe on individual privacy rights.
- **Bias**: Addressing potential biases in AI algorithms that can lead to unfair treatment or discrimination.
- **Transparency**: Making the workings of AI systems understandable to users, especially when decisions made by these systems can have serious implications.
- **Accountability**: Determining who is responsible for the actions of AI systems, particularly when they operate autonomously.

important to integrate ethical considerations into the design and deployment of these systems to prevent misuse and ensure they benefit society as a whole

Challenges of tinyML & AI: Regulations and governance

1. Connectivity

- most applications rely on a cloud-centric architecture (nodes send data to the cloud for ML inference)
- enable local processing without heavy reliance on the cloud

2. Energy consumption

- wireless communication consumes significant energy (real-time data transmission)
- edge intelligence allows for energy-efficient local processing, reducing the energy burden on devices

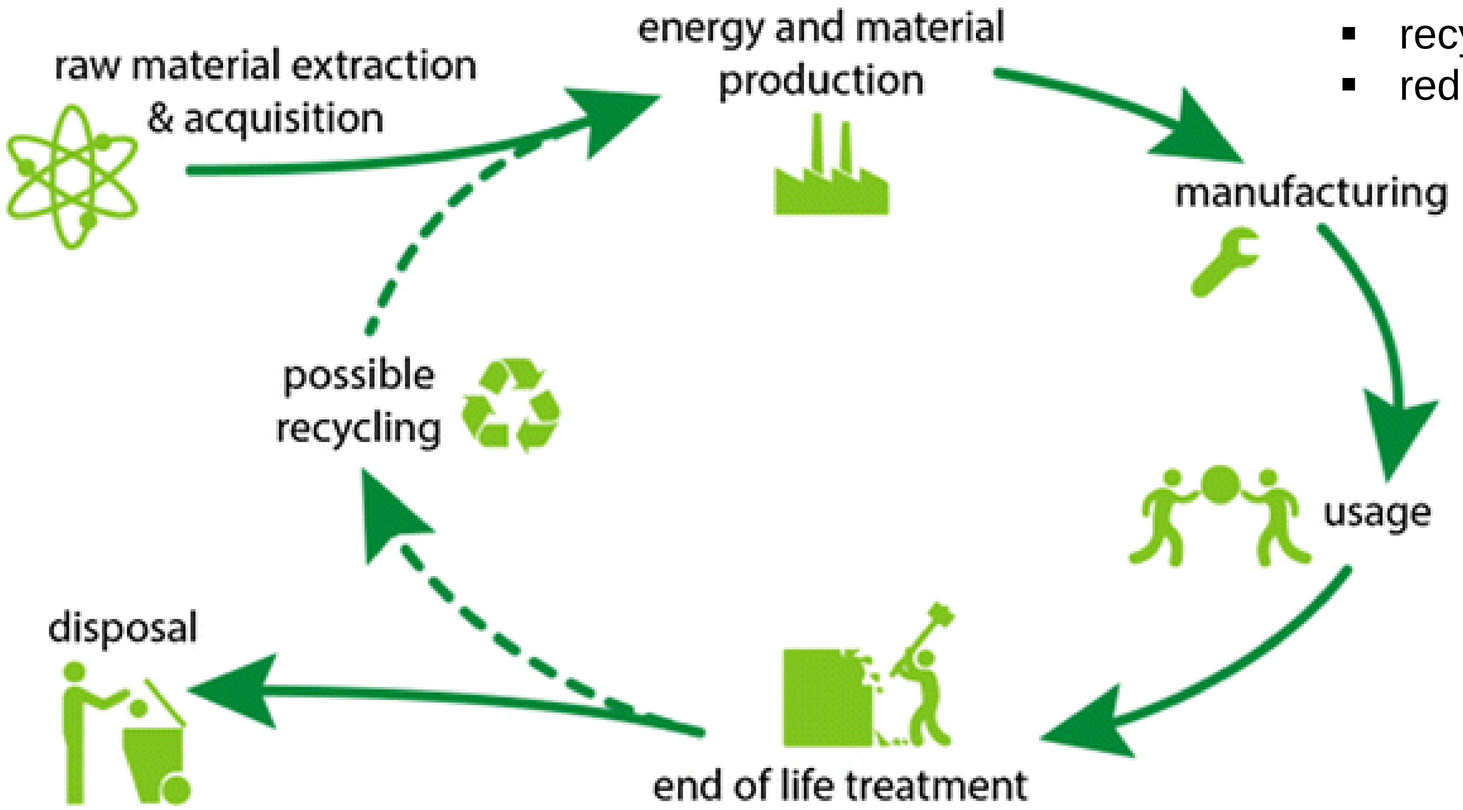
3. Regulatory landscape

- principles related to privacy, accountability, transparency, fairness, and human control are essential for effective governance

4. Standardization & security

- Developing standardized architectures for TinyML and AI is critical
- Ensuring secure deployment and addressing potential vulnerabilities

Sustainable electronics



- made with no toxic chemicals
- recyclable parts
- reduced carbon emissions

- Optimized use of energy
- Design to last

- End of life potential
- Repaired and reused

Sustainable electronics

- **No toxic chemicals**: made without toxic chemicals (incl. avoiding hazardous substances during production)

- **Recyclable parts**: products designed with components that can be easily recycled or repurposed at the end of their life cycle. (reducing electronic waste and promotes circularity)

- **Reduced carbon emissions**: minimize carbon emissions during manufacturing and in use

- **Positive impact**: benefit human health, the planet, and society (help prevent harmful chemicals from entering the environment and reduce the overall ecological footprint of electronic devices)

Challenges of tinyML & AI: Sustainability

Sustainable – climate, energy efficiency, waste, etc

Sustainability means meeting our own needs without compromising the ability of future generations to meet their own needs

Key considerations;

- **Energy efficiency**: operate for extended periods on minimal power (shrink AI/ML models without sacrificing performance)

- **Resource constraints**: ensuring devices operate efficiently while minimizing their impact on resources (run on resource-constrained devices)

- **Environmental impact**: Balancing the benefits of AI with its potential carbon footprint and waste generation is a critical

- **Democratization and local innovation**: offers a sustainable path toward democratizing AI, fostering local innovation, and addressing regional challenges

- **SDGs and connectivity**

Challenges of tinyML & AI: e-waste

Electronic waste describes discarded electrical or electronic devices

32%

62 M
Metric Tones
produced in 2022

82%
from 2010

82 M
Metric Tones
in 2030

- e-waste generated by **outdated** or **malfunctioning** TinyML devices
- rapid pace of tech. advancements leads to more frequent **device upgrades** and **replacements**
- e-waste **recycling** activities occur using environmentally unsound techniques
- inadequate **regulations**, recycling **infrastructure**, and **training**

Overcoming some of the tinyML Challenges

1. Processing capacity optimization

- TinyML devices operate with limited computational resources
- Balancing accuracy and computational efficiency ensure efficient execution of ML models

2. Energy consumption

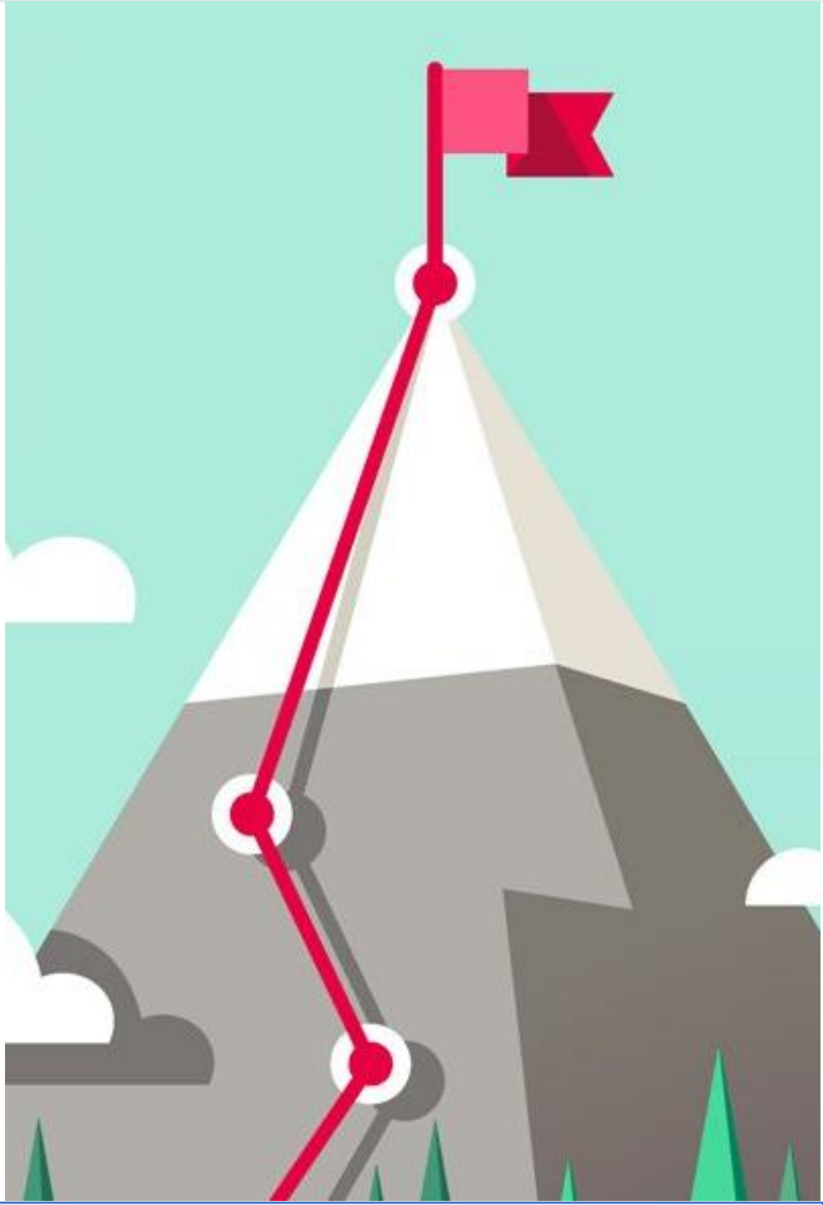
- Energy-efficient local processing, reducing the energy burden on devices

3. Improved reliability

- Ensuring robustness against environmental variations, sensor noise, and hardware limitations is essential

4. Maintenance of learning models' accuracy

- Drift, concept shift, and model degradation must be addressed to sustain reliable performance



ITU Events

AI for Good
Global Summit

Accelerating the United Nations Sustainable Development Goals

30-31 May 2024

Geneva, Switzerland

aiforgood.itu.int/

40 UN PARTNERS

ITU

ITU Events

AI for Good Workshops

tinyML: Pioneering sustainable solutions in resource-limited environments

Thursday, 30 May 2024

09:30 - 17:30 Geneva (CEST)

aiforgood.itu.int

40 UN PARTNERS

ITU