

IESTI01 - TinyML

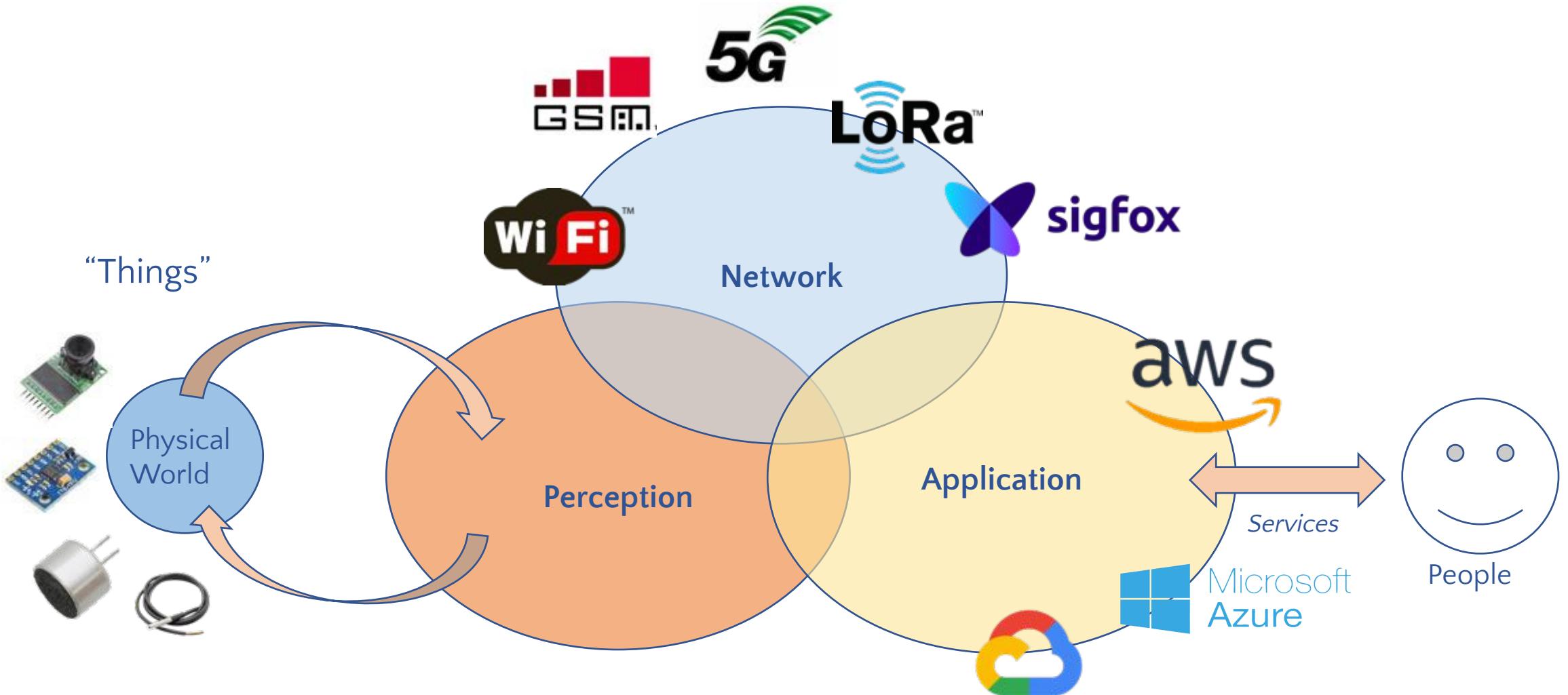
Responsible AI &
Course Wrapup

Prof. Marcelo Rovai

July 28th, 2021

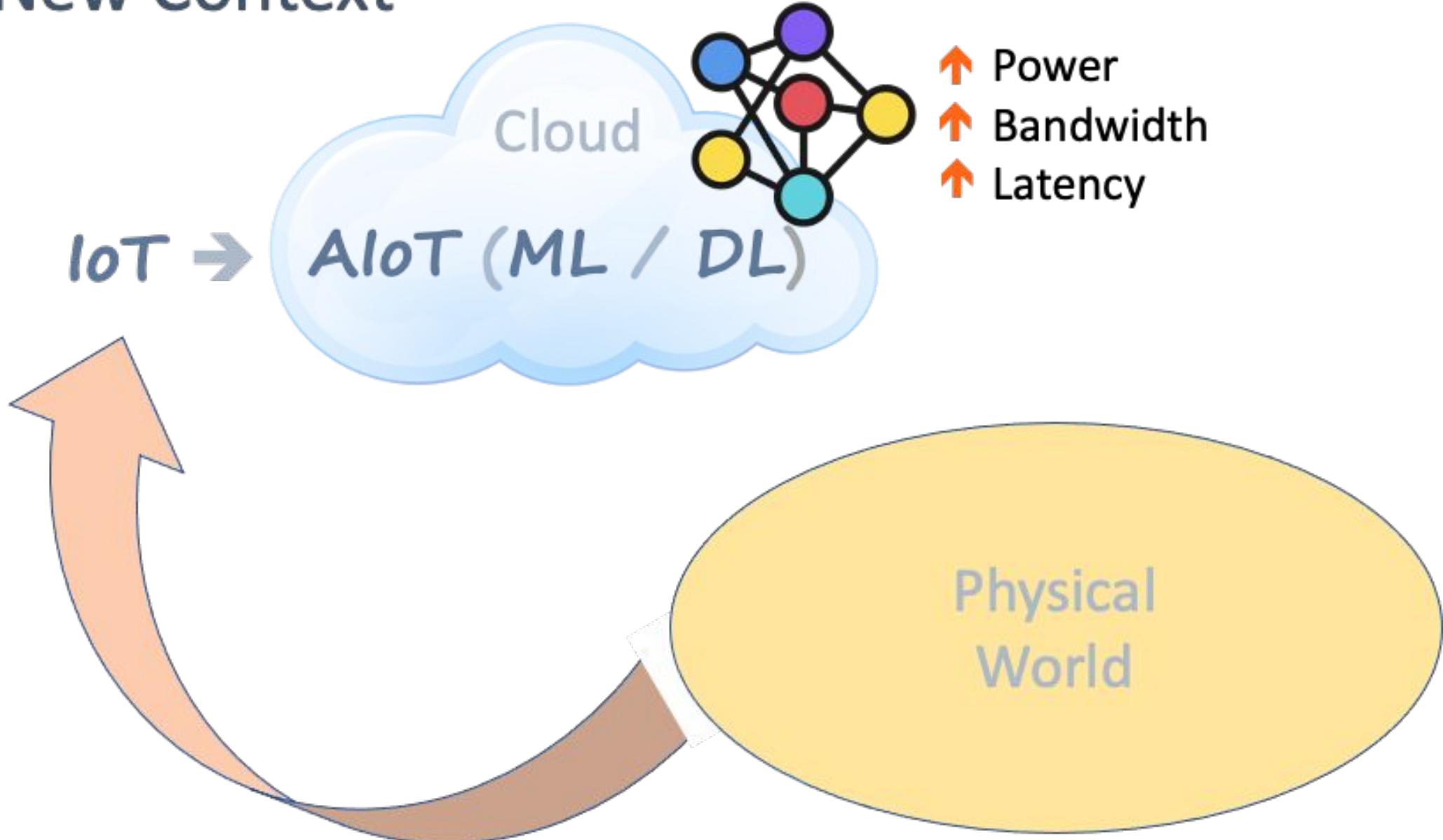


Classic IoT Architecture

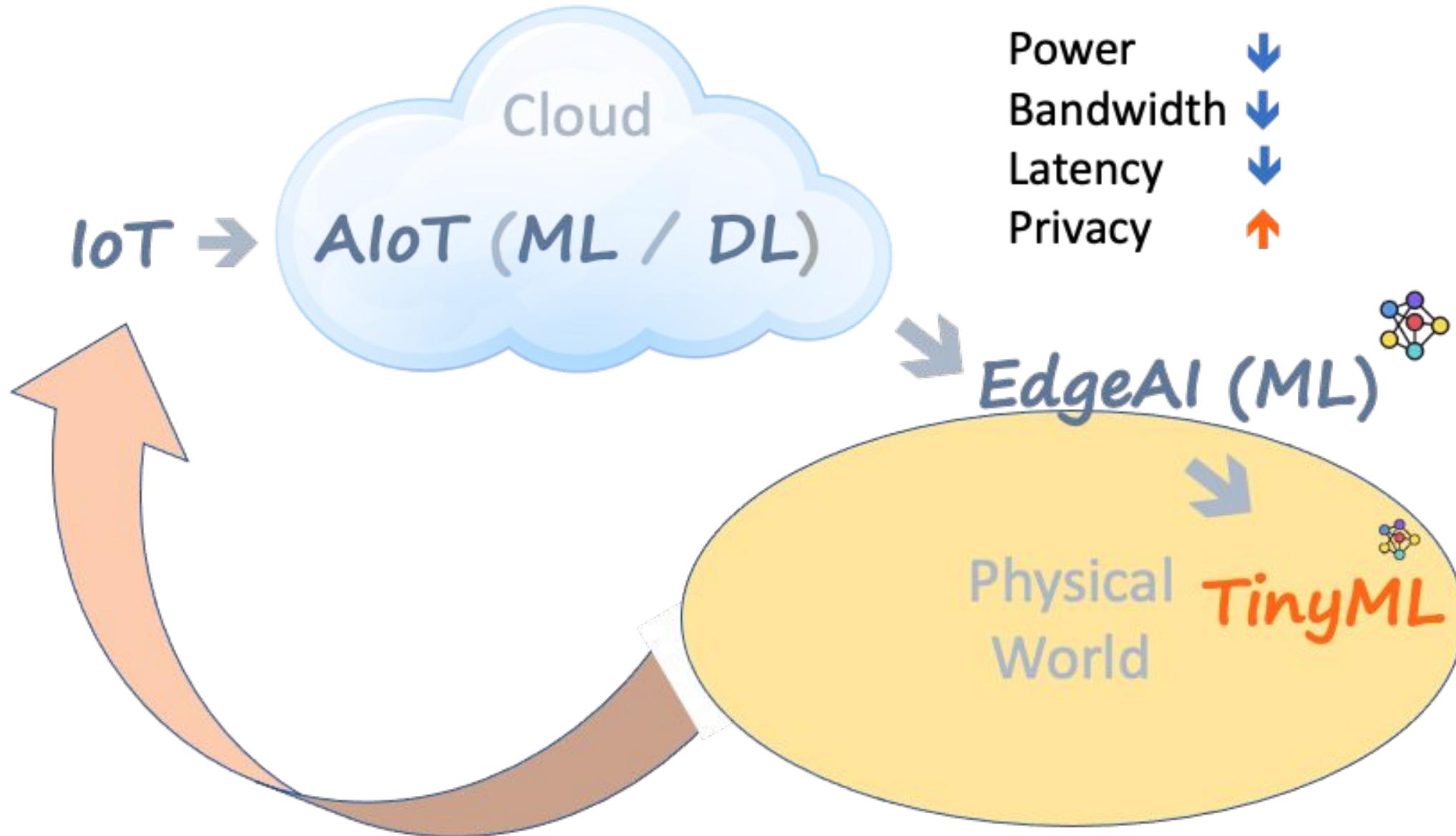


5 Quintillion bytes of data produced every day by IoT, but less than 1% is used. HBR/CISCO

New Context



New Context



Applications: EdgeML & TinyML

Autonomous Car Control



KeyWord Spotting



Environmental Control



Image Recognition



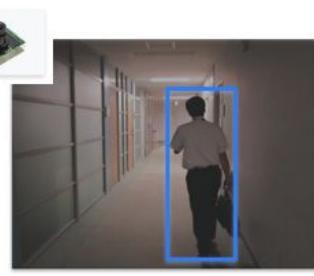
Motion & biometric



Image Spot



TinyML Application Areas



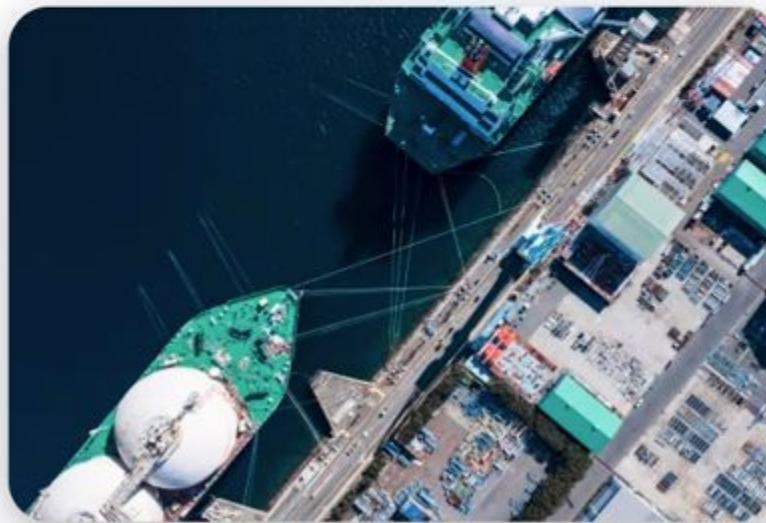
Predictive Maintenance



Motion, current, audio and camera

- Industrial
- White goods
- Infrastructure
- Automotive

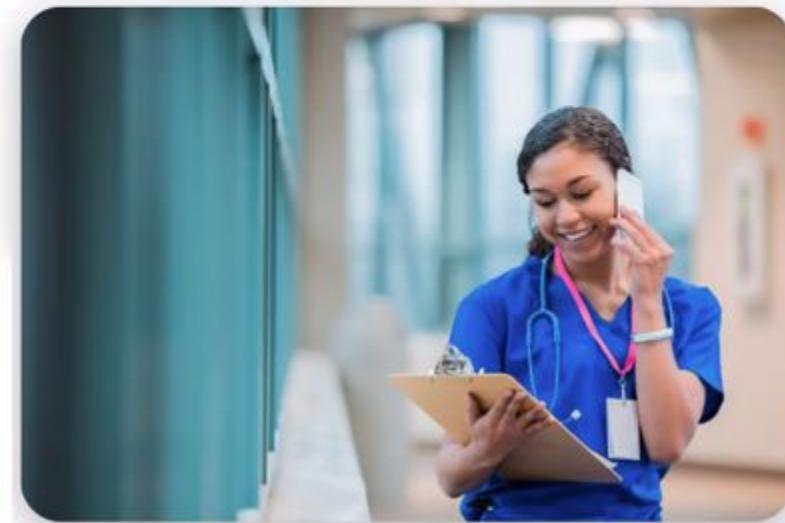
Asset Tracking & Monitoring



Motion, temp, humidity, position, audio and camera

- Logistics
- Infrastructure
- Buildings

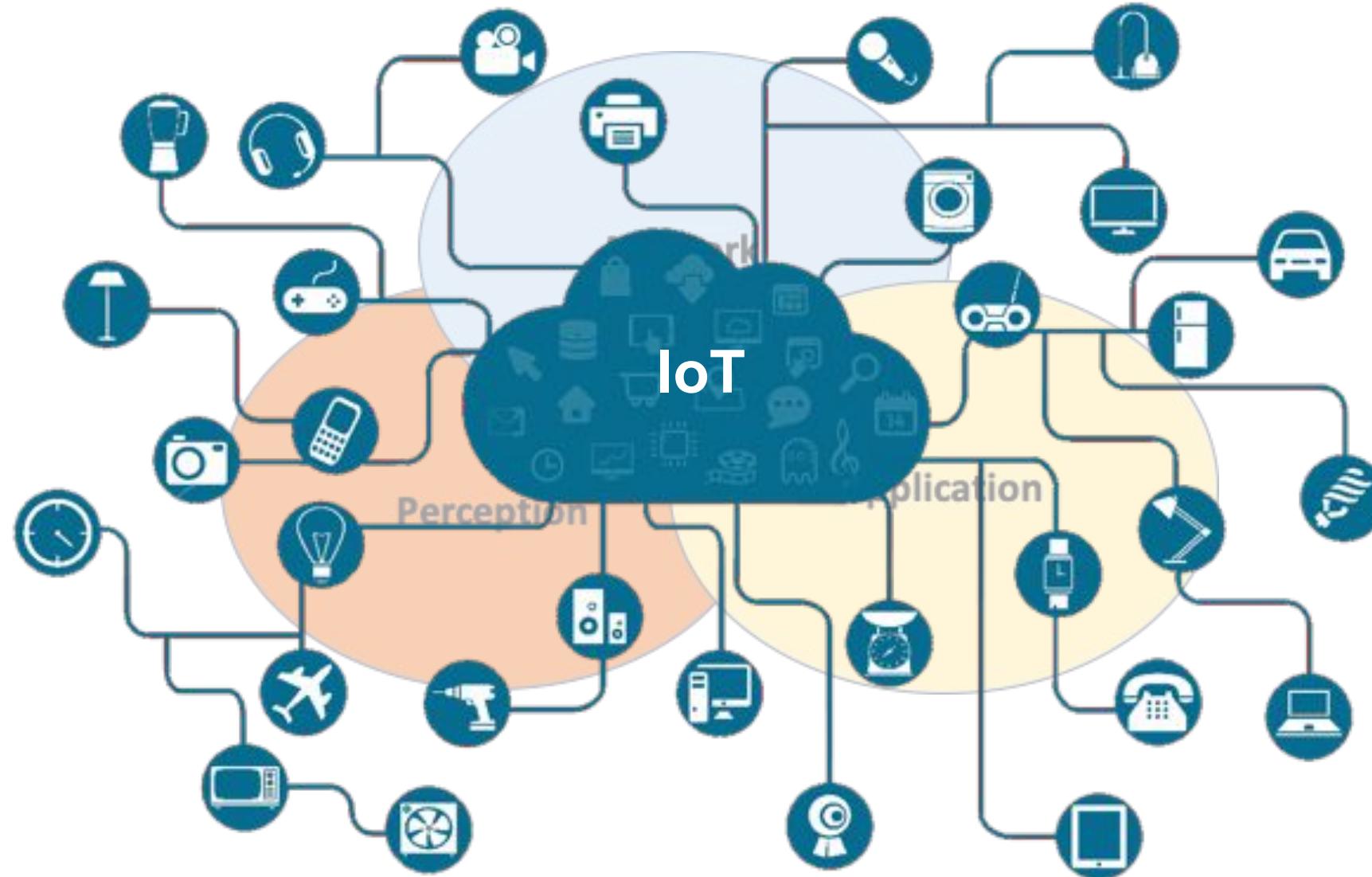
Human & Animal Sensing



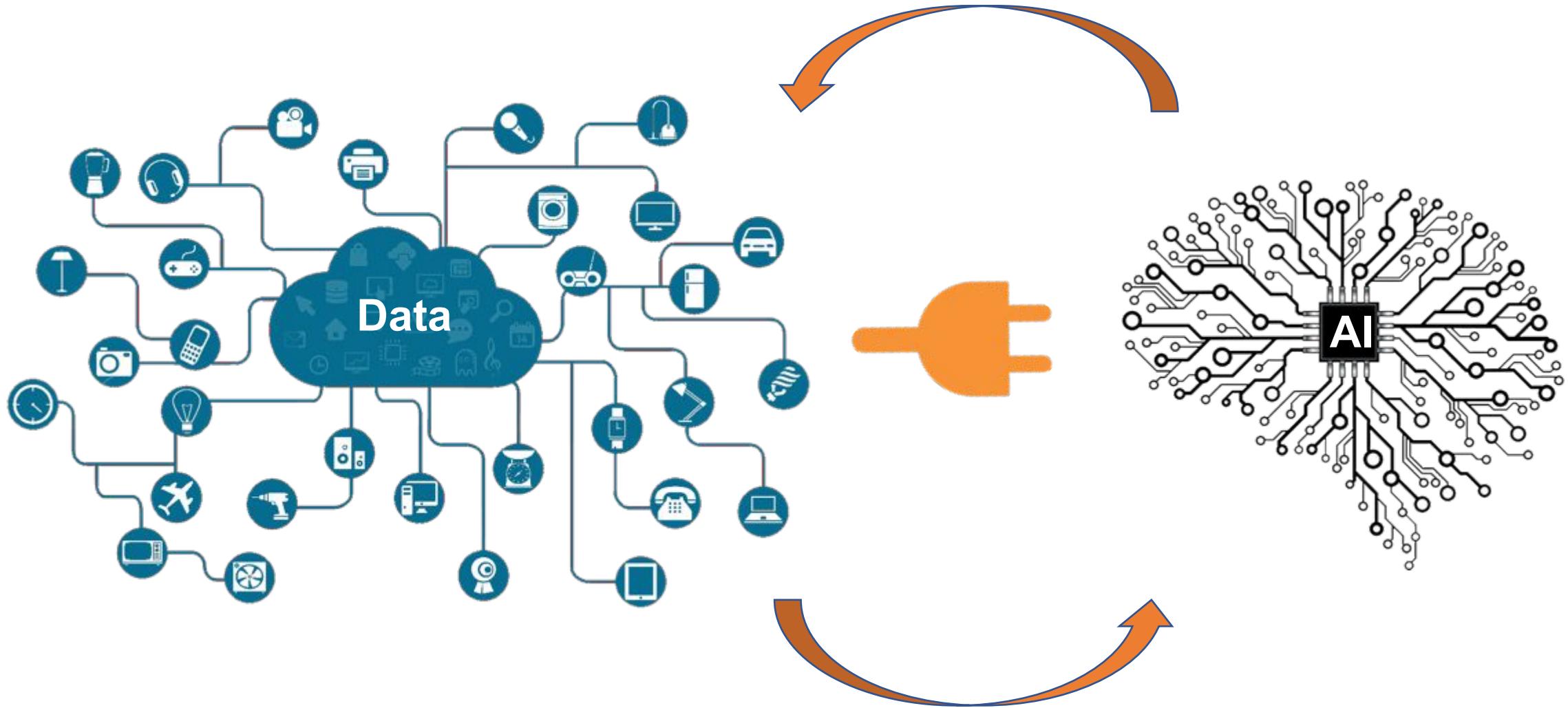
Motion, radar, audio, PPG, ECG

- Health
- Consumer
- Industrial

IoT - Architecture



Endpoints devices □ Data + AI □ Value



Endpoints Have Sensors, Tons of Sensors

Motion Sensors

Gyroscope, radar,
magnetometer, accelerator

Acoustic Sensors

Ultrasonic, Microphones,
Geophones, Vibrometers

Environmental Sensors

Temperature, Humidity,
Pressure, IR, etc.

Touchscreen Sensors

Capacitive, IR

Image Sensors

Thermal, Image

Biometric Sensors

Fingerprint, Heart rate, etc.

Force Sensors

Pressure, Strain

Rotation Sensors

Encoders

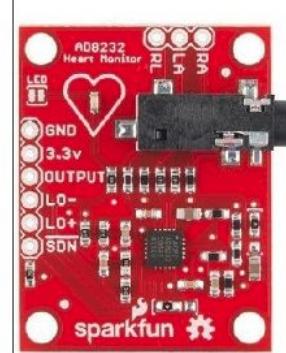
Biometric Sensors



Non-invasive Glucose Monitoring



Fingerprint + Photoplethysmography (PPG)



ECG
Sensor

Endpoints Have Sensors, Tons of Sensors

Motion Sensors

Gyroscope, radar,
magnetometer, accelerator

Acoustic Sensors

Ultrasonic, Microphones,
Geophones, Vibrometers

Environmental Sensors

Temperature, Humidity,
Pressure, IR, etc.

Touchscreen Sensors

Capacitive, IR

Image Sensors

Thermal, Image

Biometric Sensors

Fingerprint, Heart rate, etc.

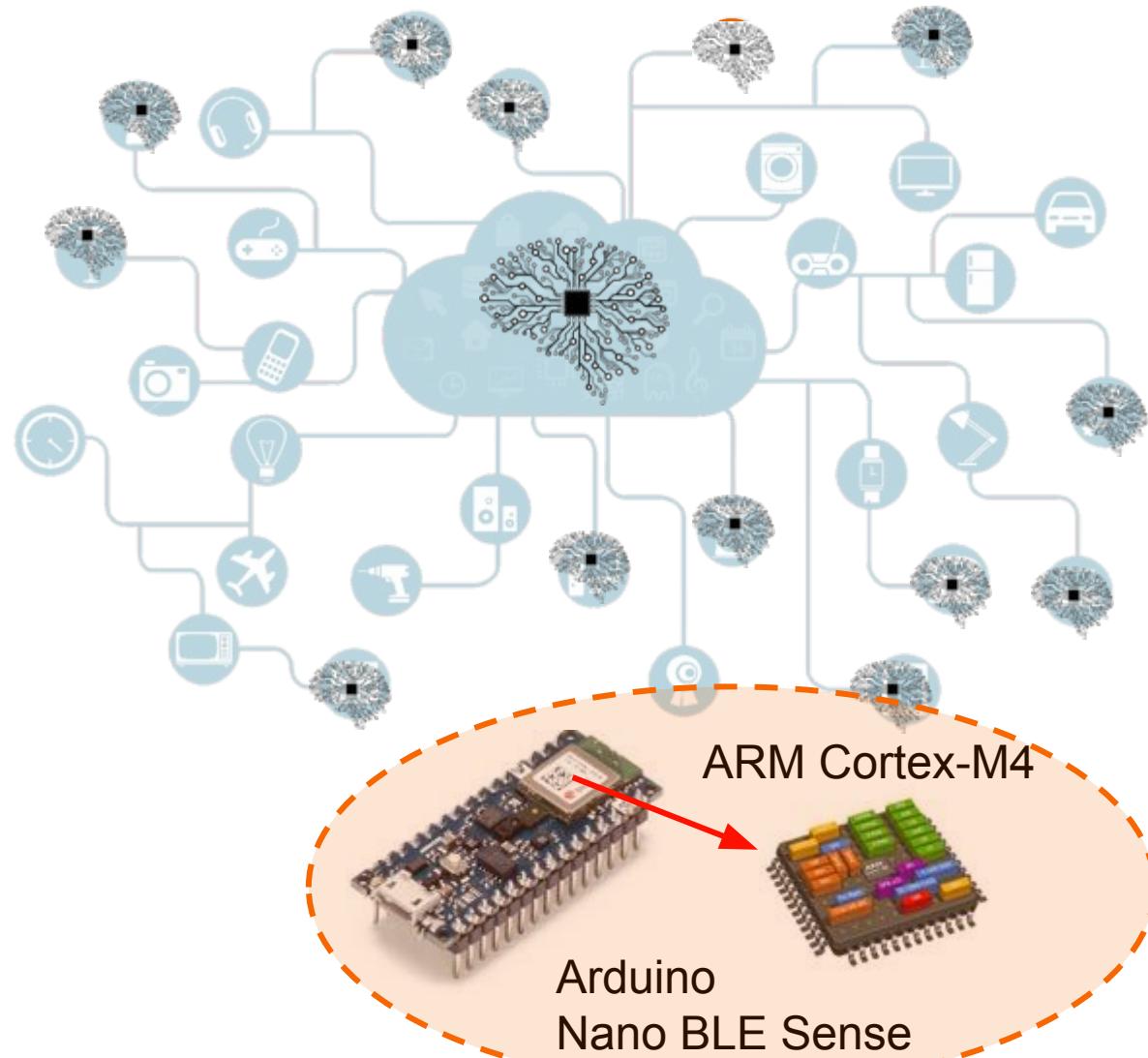
Force Sensors

Pressure, Strain

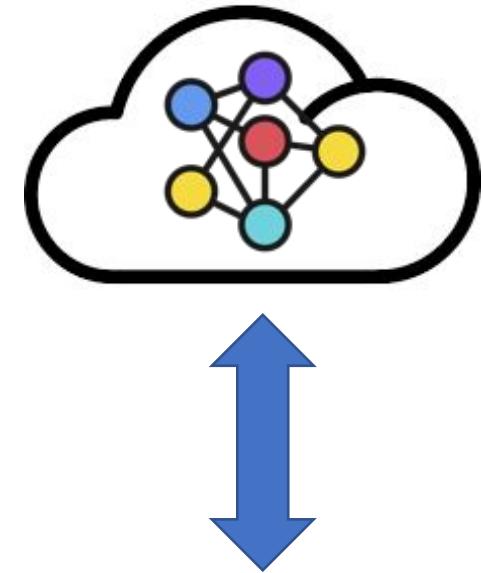
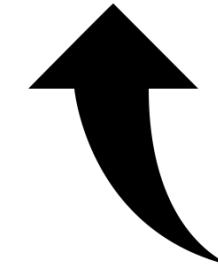
Rotation Sensors

Encoders

ML (AI) at the “edge of the edge” - TinyML

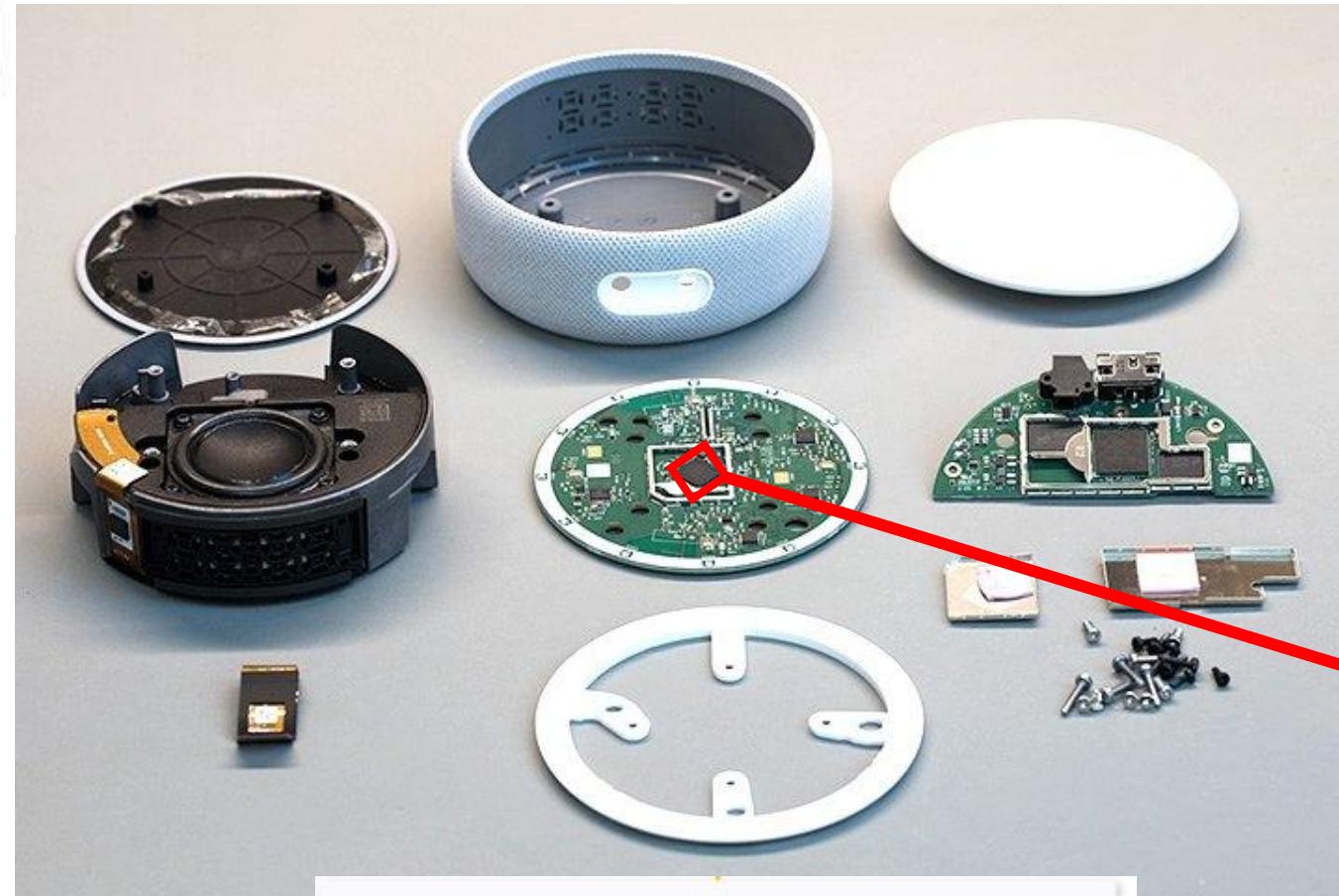


TinyML enables machine intelligence right next to the physical world



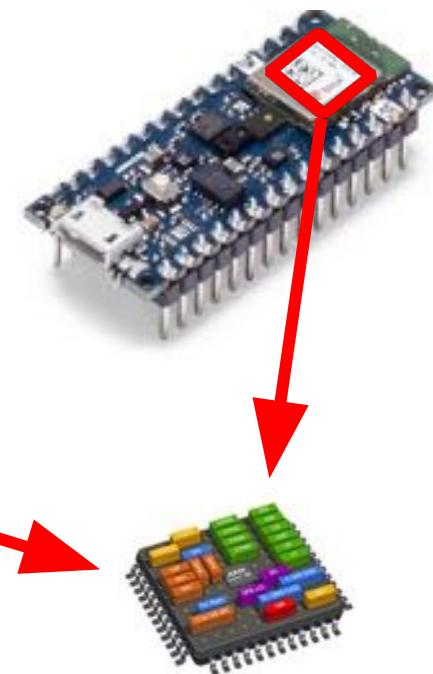
ML at microprocessor level avoid issues as Latency, Power Consuming, and Security

Echo-Dot Teardown vs Arduino Nano BLE Sense



MediaTek 7658CSN: Wi-Fi +ARM® Cortex-R4

Nordic nRF52840-M4

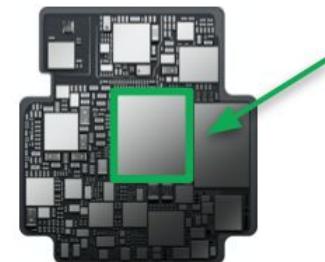


MCUs enable **TinyML**

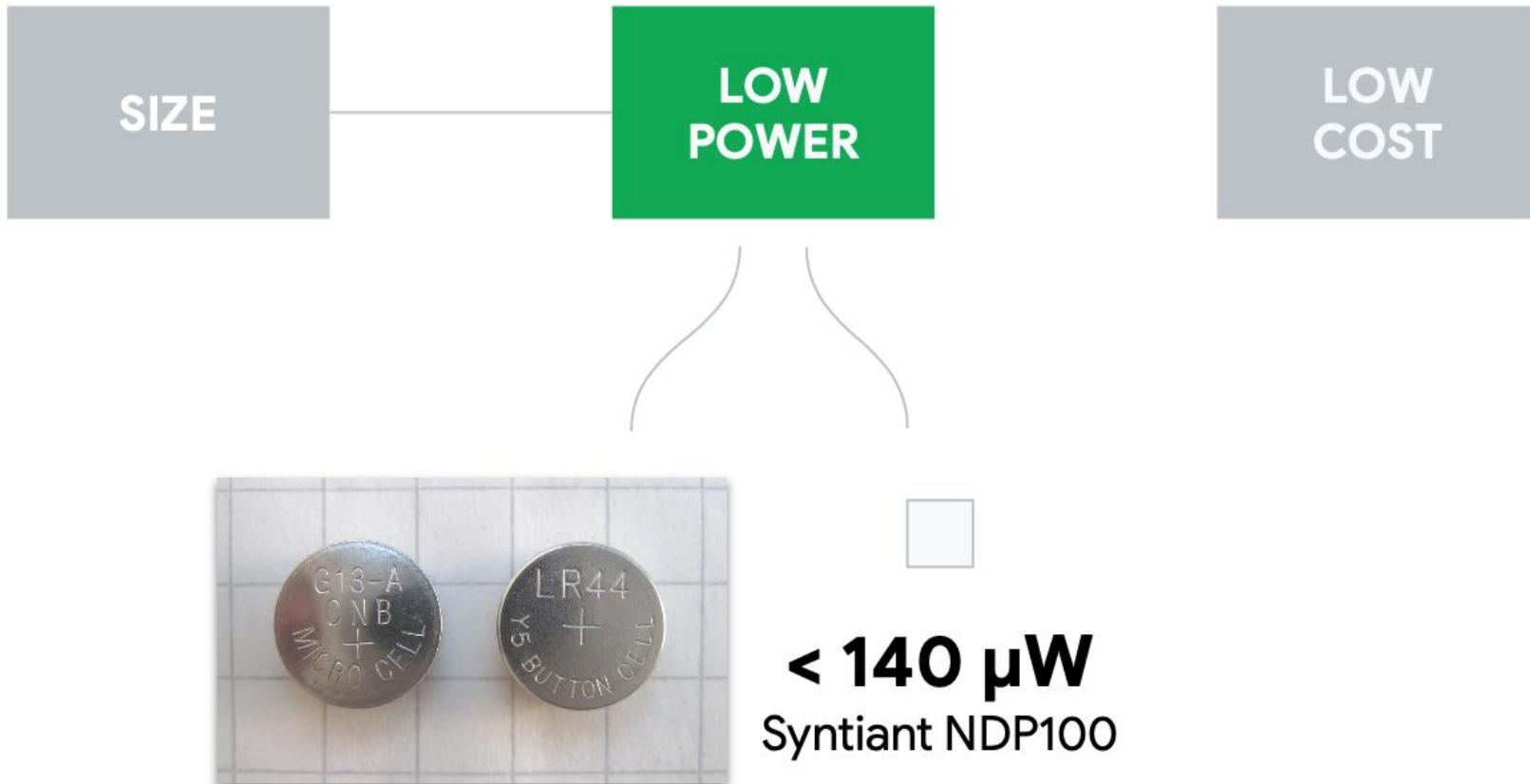
SIZE

LOW
POWER

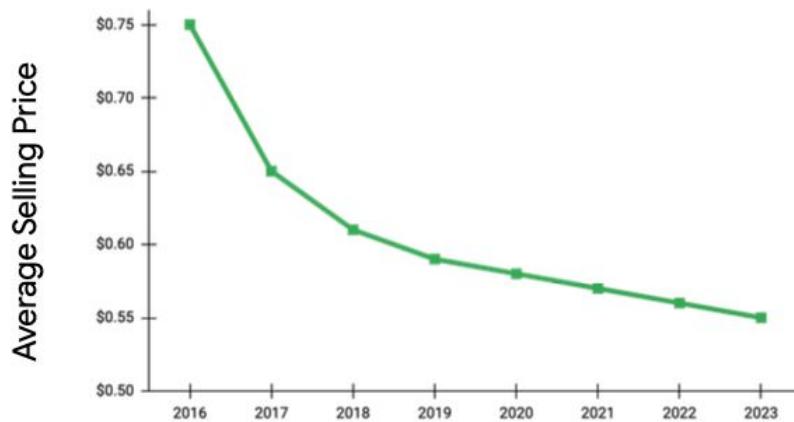
LOW
COST



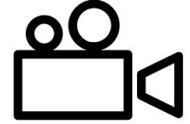
MCUs enable **TinyML**



MCUs enable **TinyML**



Hardware



Anomaly Detection
Sensor Classification
20 KB



Rpi-Pico
(Cortex-M0+)



Arduino Nano
(Cortex-M4)



Arduino Pro
(Cortex-M7)

Source: Edge Impulse

EdgeML

TinyML

KeyWord Spotting
Audio Classification
50 KB

Image
Classification
250 KB+



Video
Classification
2 MB+



Object Detection
Complex Voice
Processing
1 MB+

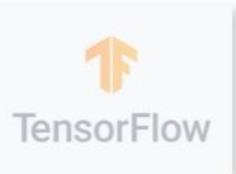


Raspberry Pi
(Cortex-A)

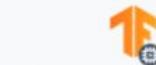


Jetson Nano
(Cortex-A + GPU)

Software



TensorFlow



TensorFlow Lite

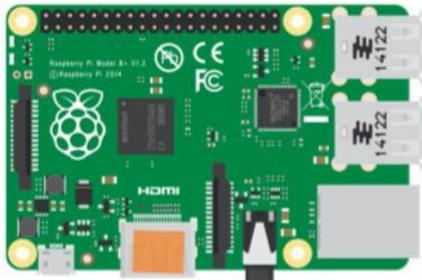
Train a model

Convert
model

Optimize
model

Deploy
model at
Edge

Make
inferences
at Edge



Raspberry Pi



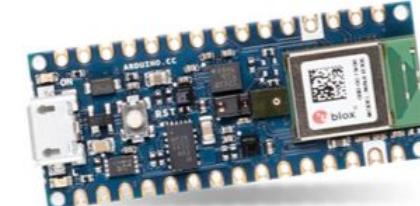
Linux



iOS

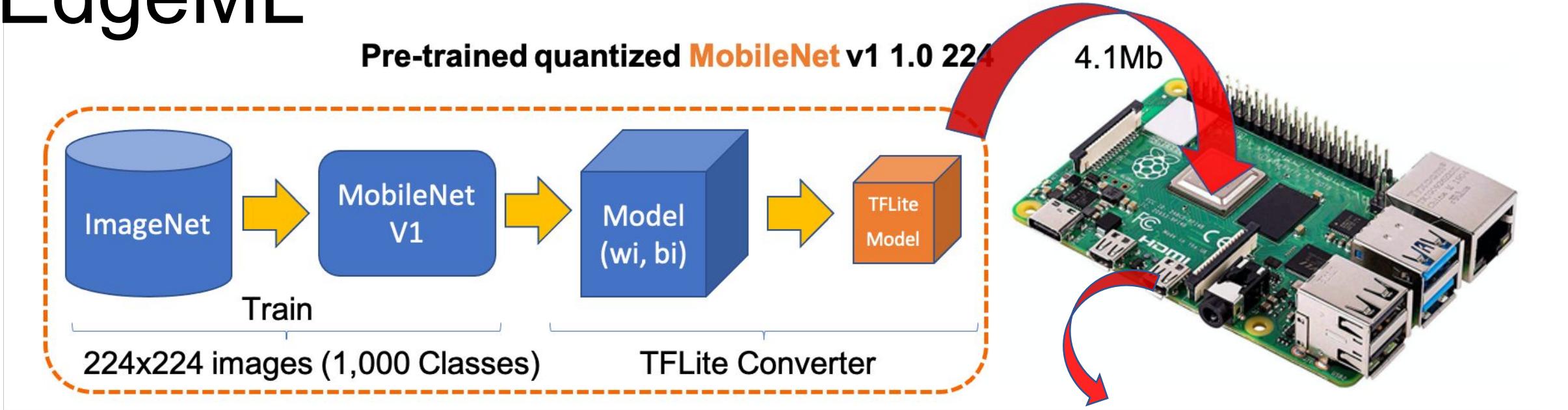


(TFL Micro)



Microcontroller

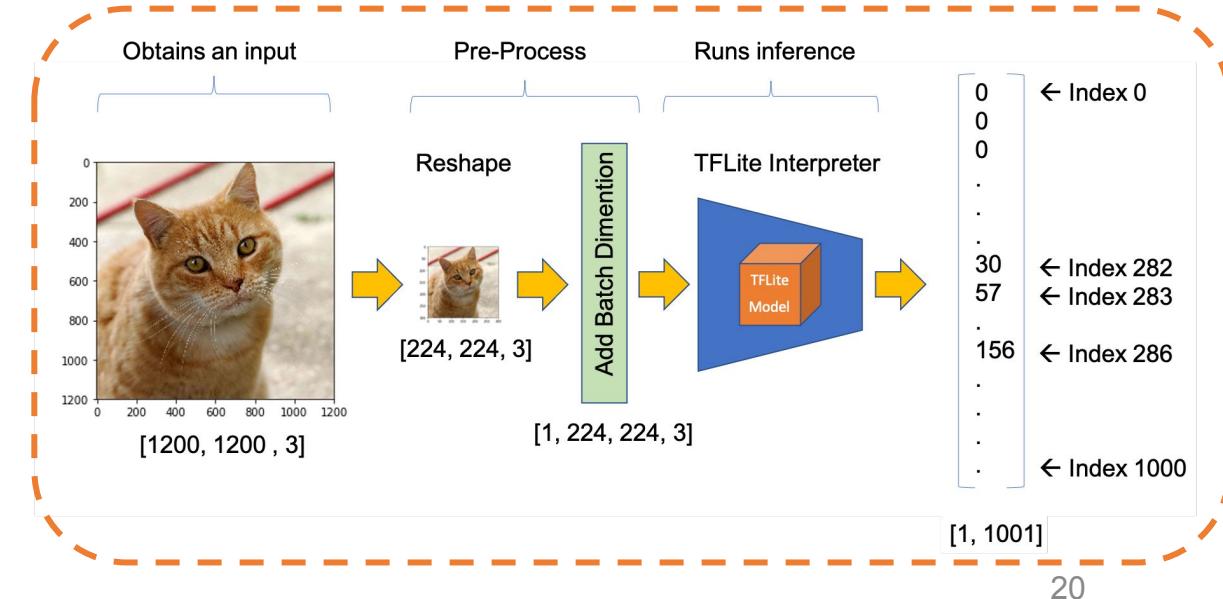
EdgeML



Exploring IA at the Edge!



UNIFEI TELC03 30-Oct-20

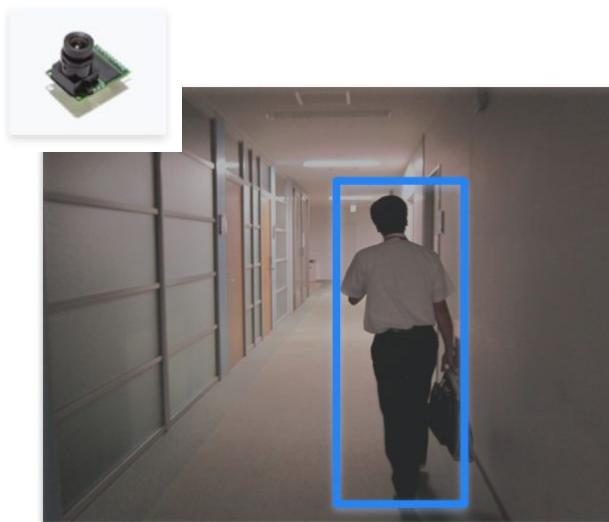


TinyML Application Examples

Sound



Vision



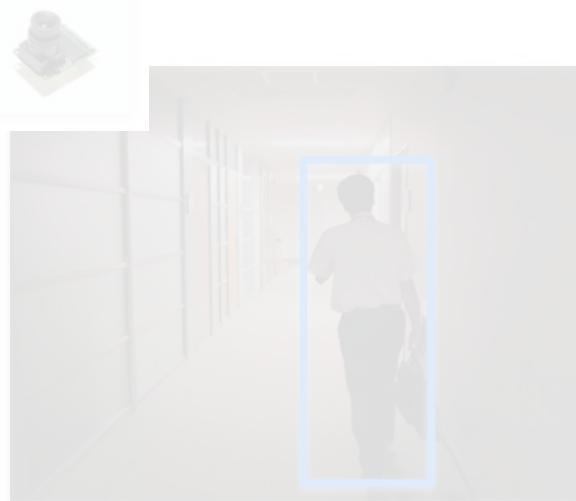
Vibration



Sound



Vision



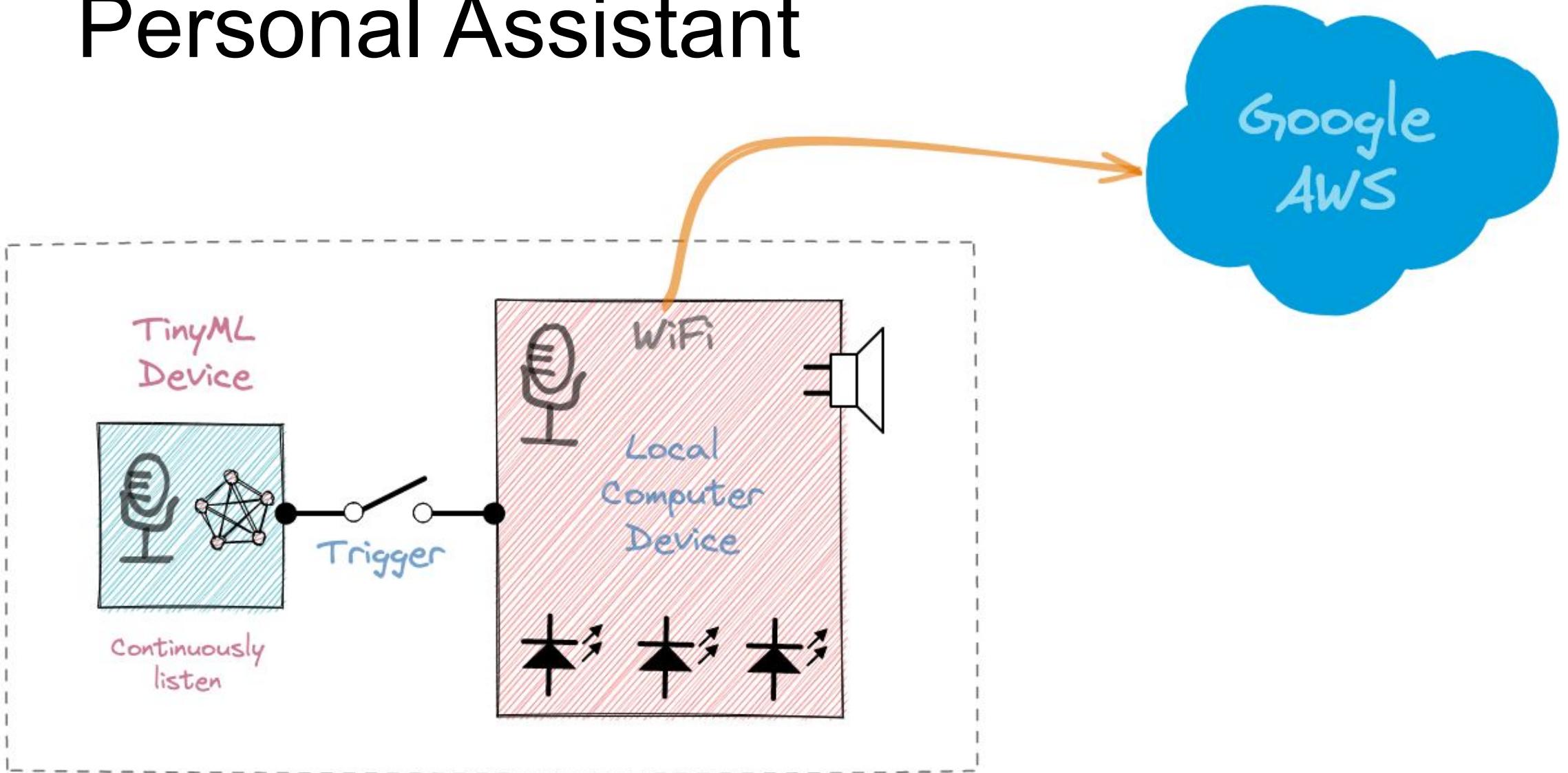
Vibration



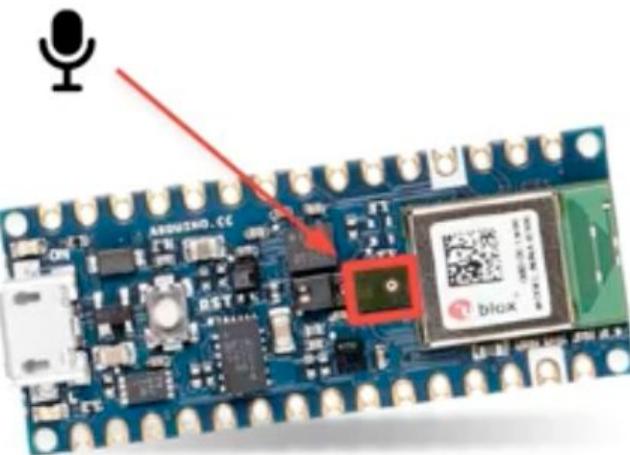
Personal Assistant



Personal Assistant



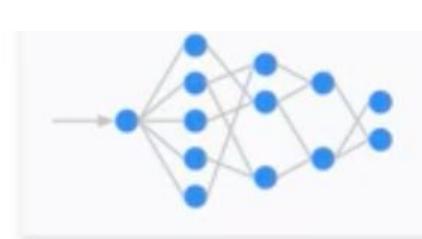
“Cascade” Detection: multi-stage model



- 1 Continuously listen on the microcontroller

2

- Process the data with **TinyML** at the edge



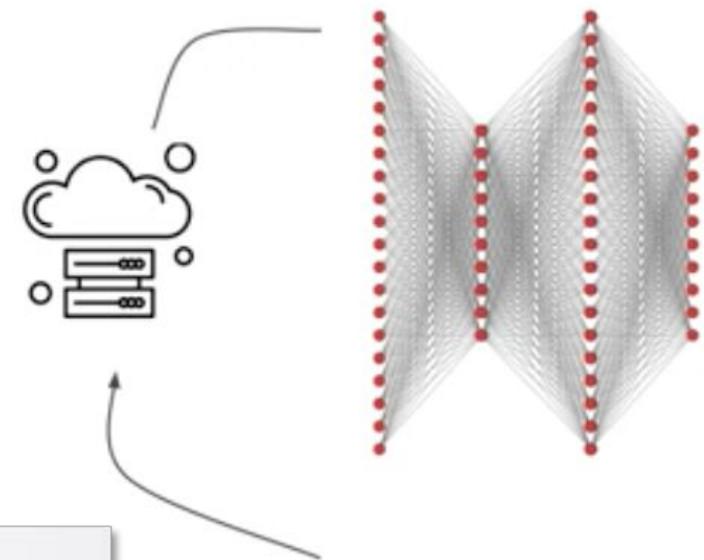
3

- Process on a secondary larger model on a larger local device



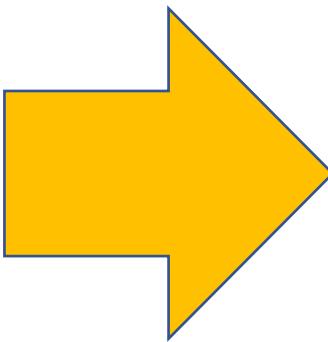
5

- Process the full speech data with a large model in the cloud

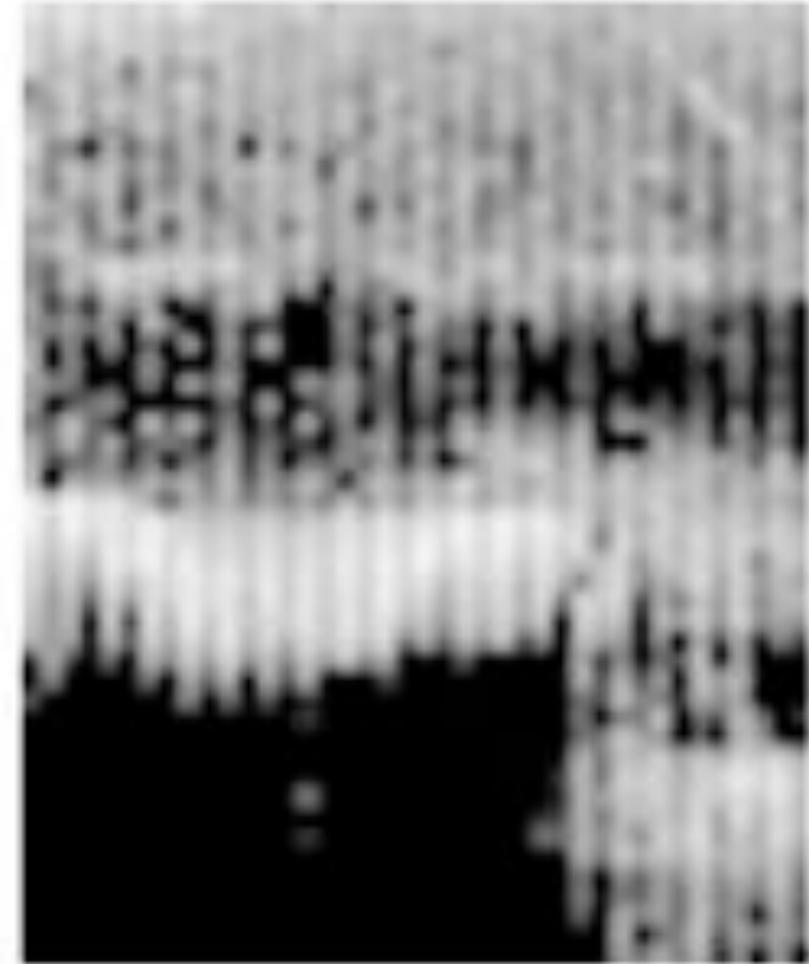


4

- Send the data to the cloud when triggered

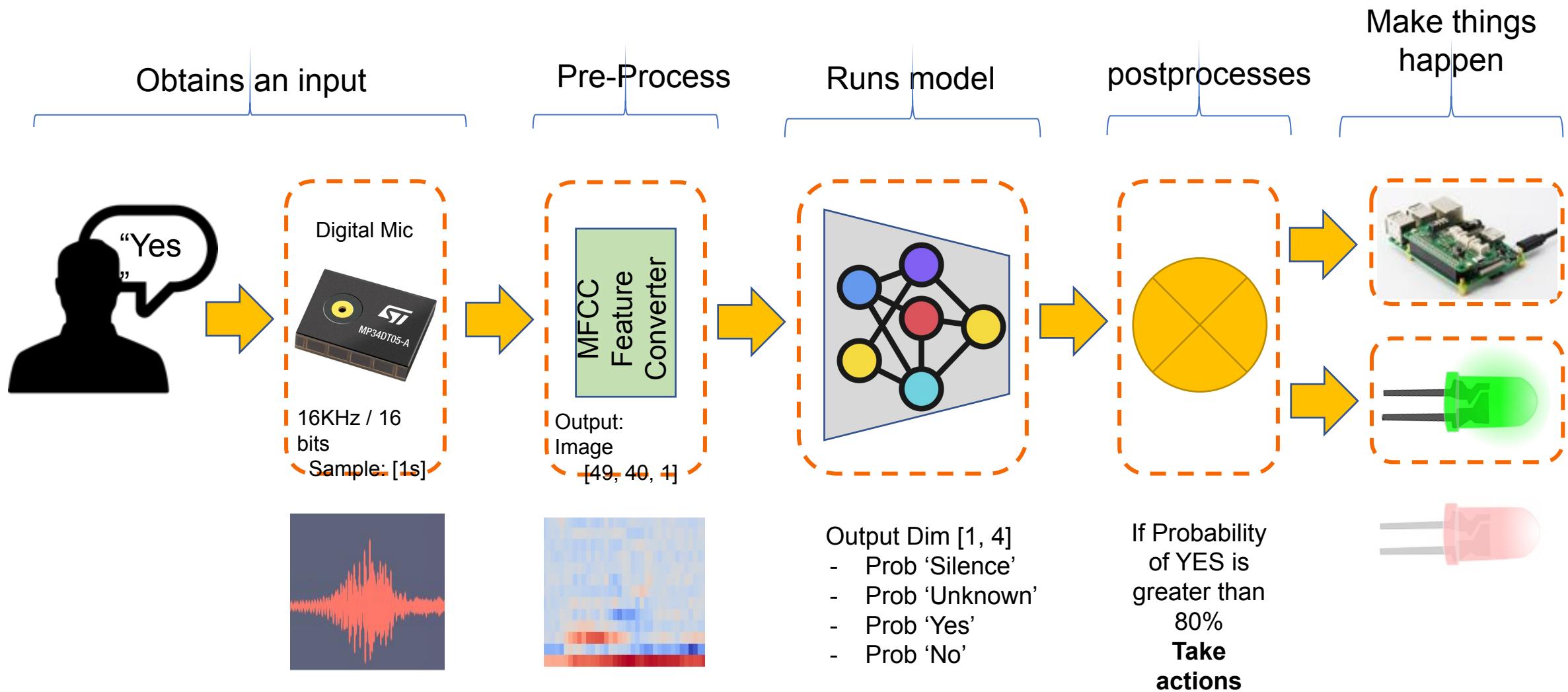


Sound

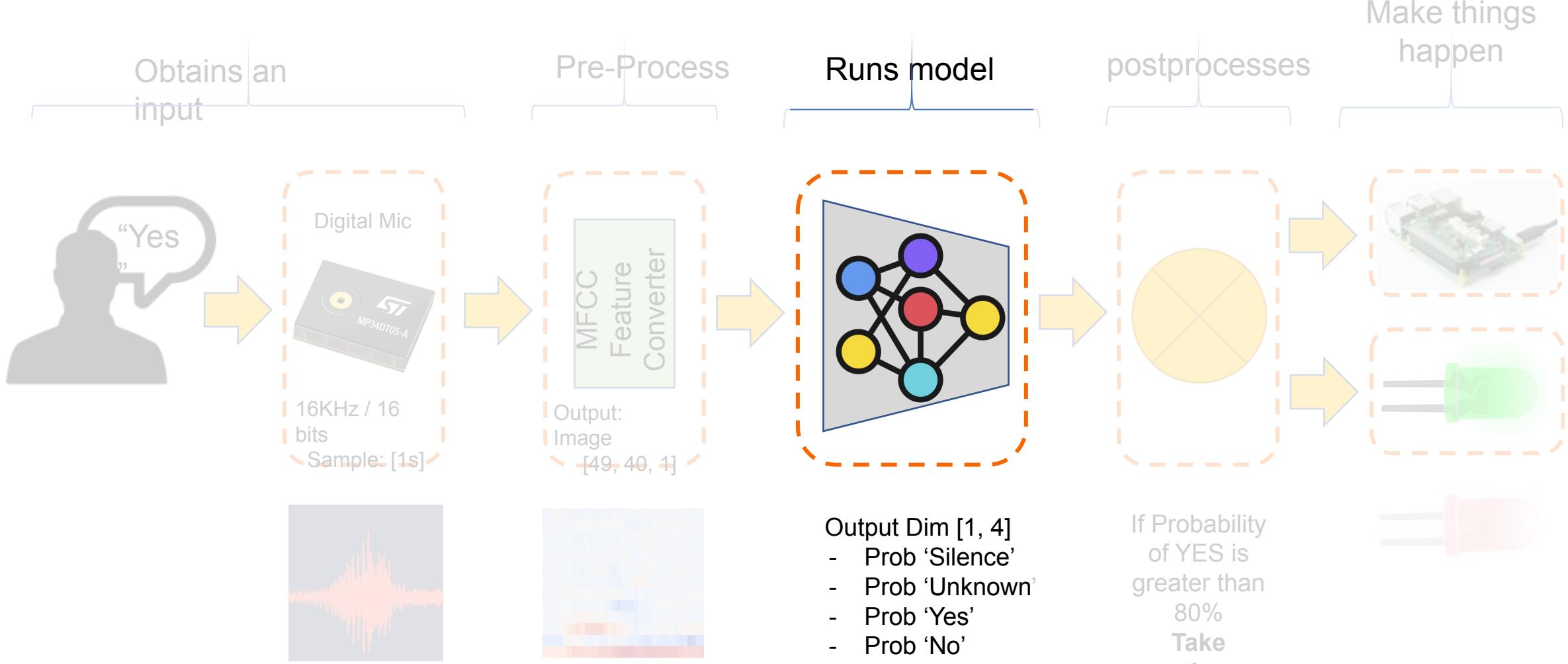


Image

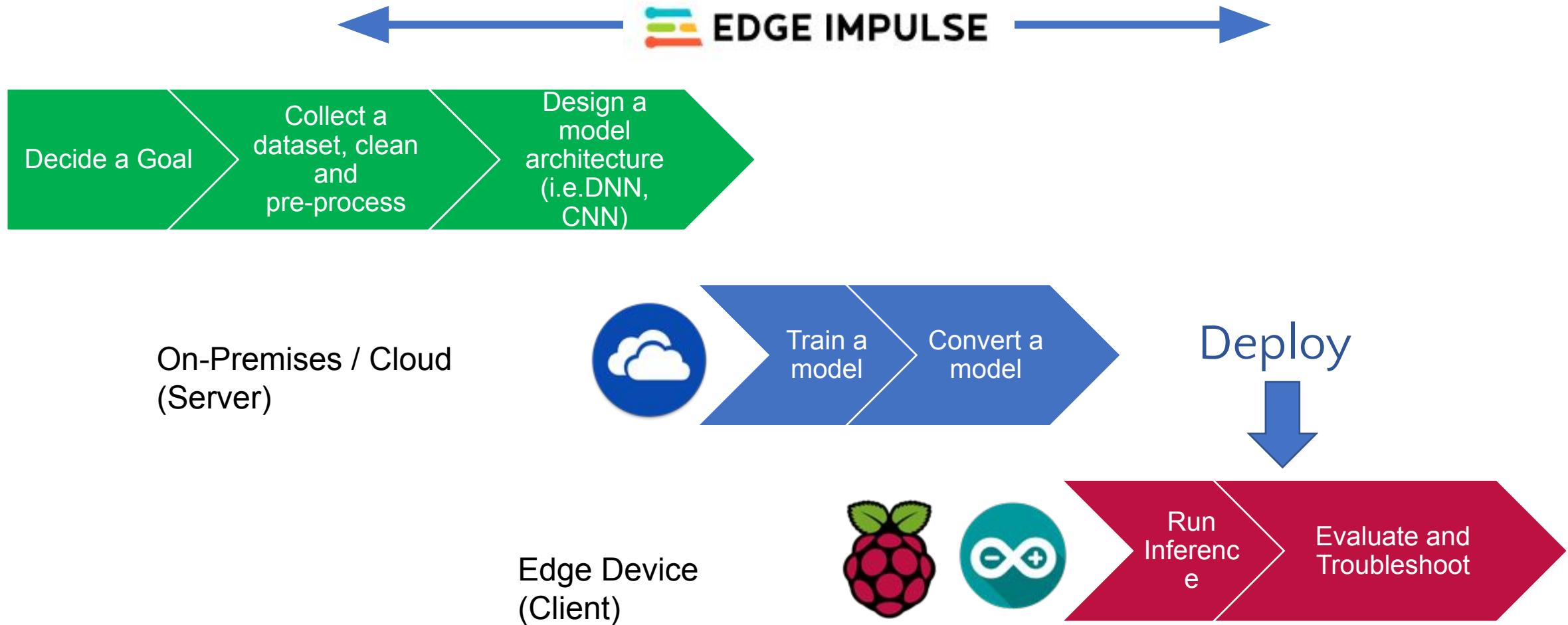
KeyWord Spotting (KWS) - Inference



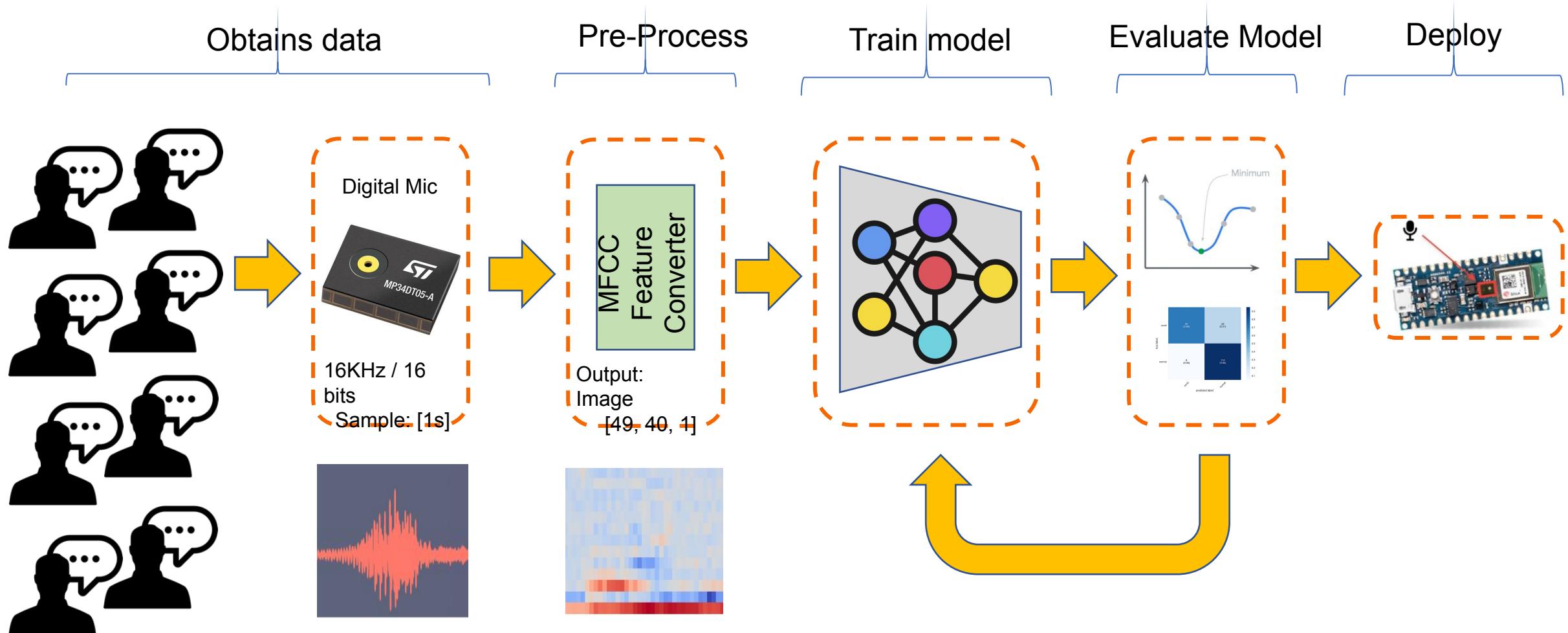
KeyWord Spotting (KWS) - Model



Machine Learning Workflow



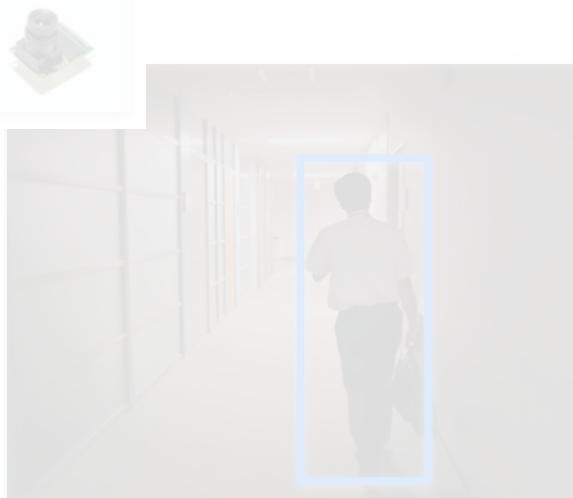
KeyWord Spotting (KWS) – Create Model (Training)



Sound



Vision



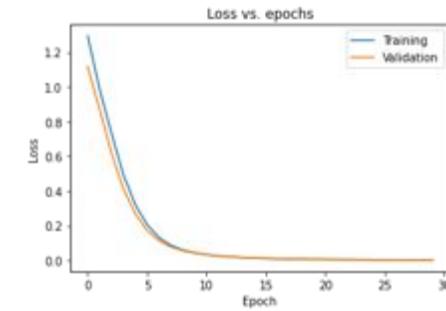
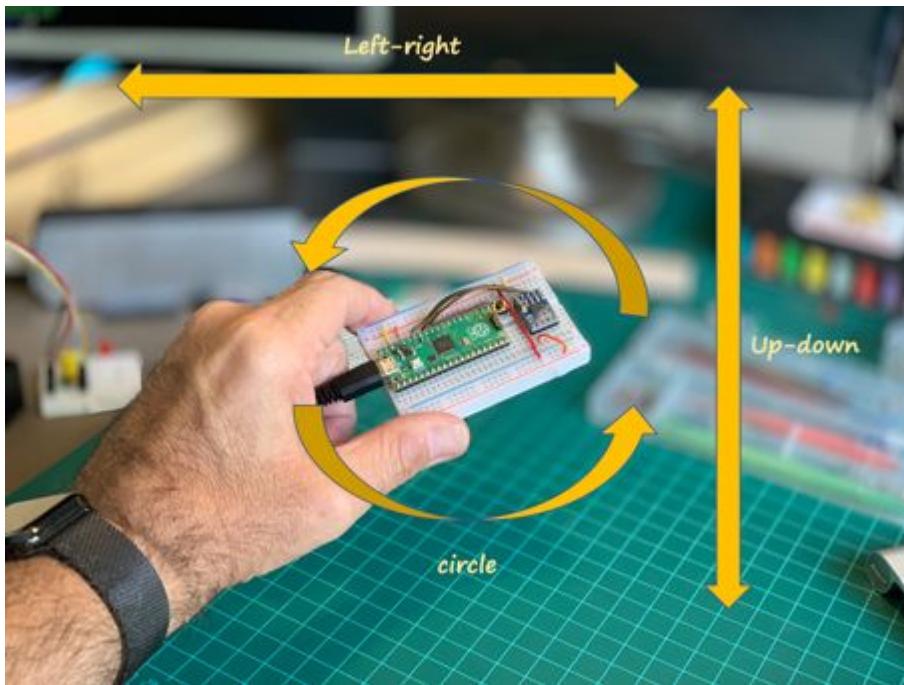
Vibration

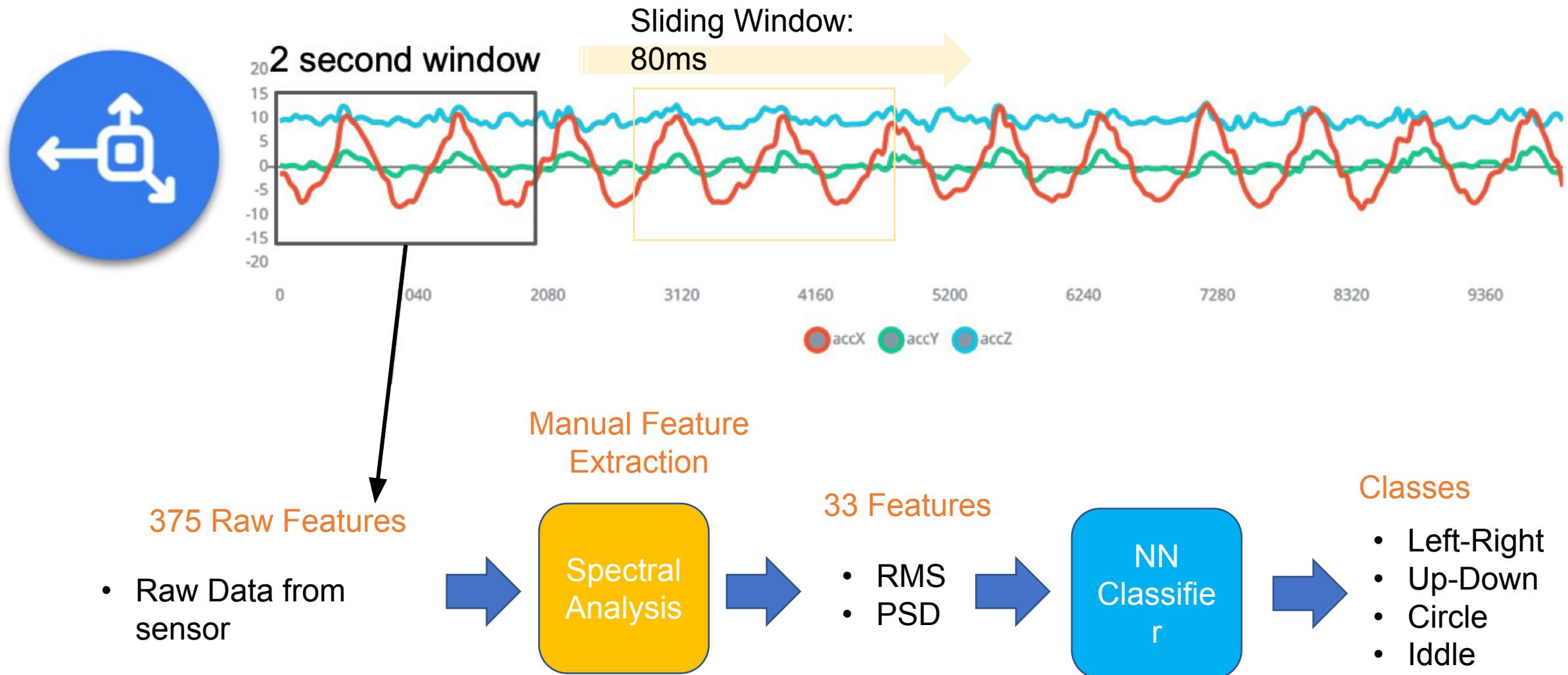


Gesture Classification - Rpi-Pico

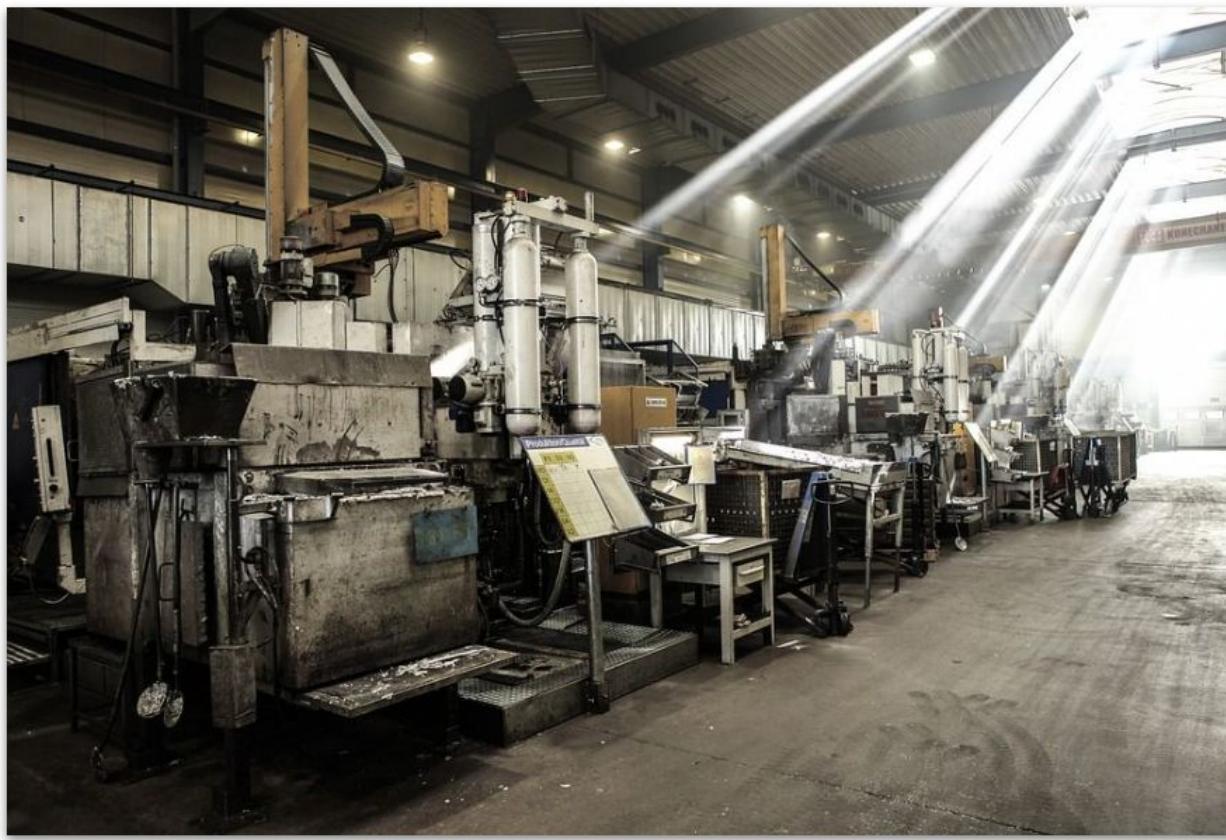


fritzing

A screenshot of a terminal window titled '/dev/cu.usbmodem0000000000000001'. The window shows three sets of text output, each starting with 'Sampling...' followed by 'Predictions' and a timestamp. The predictions are for four classes: 'circle', 'left_right', 'resting', and 'up_down'. Arrows point from the text to the corresponding gesture icons in the image above. The bottom of the window includes standard terminal controls like 'Autoscroll', 'Show timestamp', 'Both NL & CR', '115200 baud', and 'Clear output'.



Application: Factory machinery

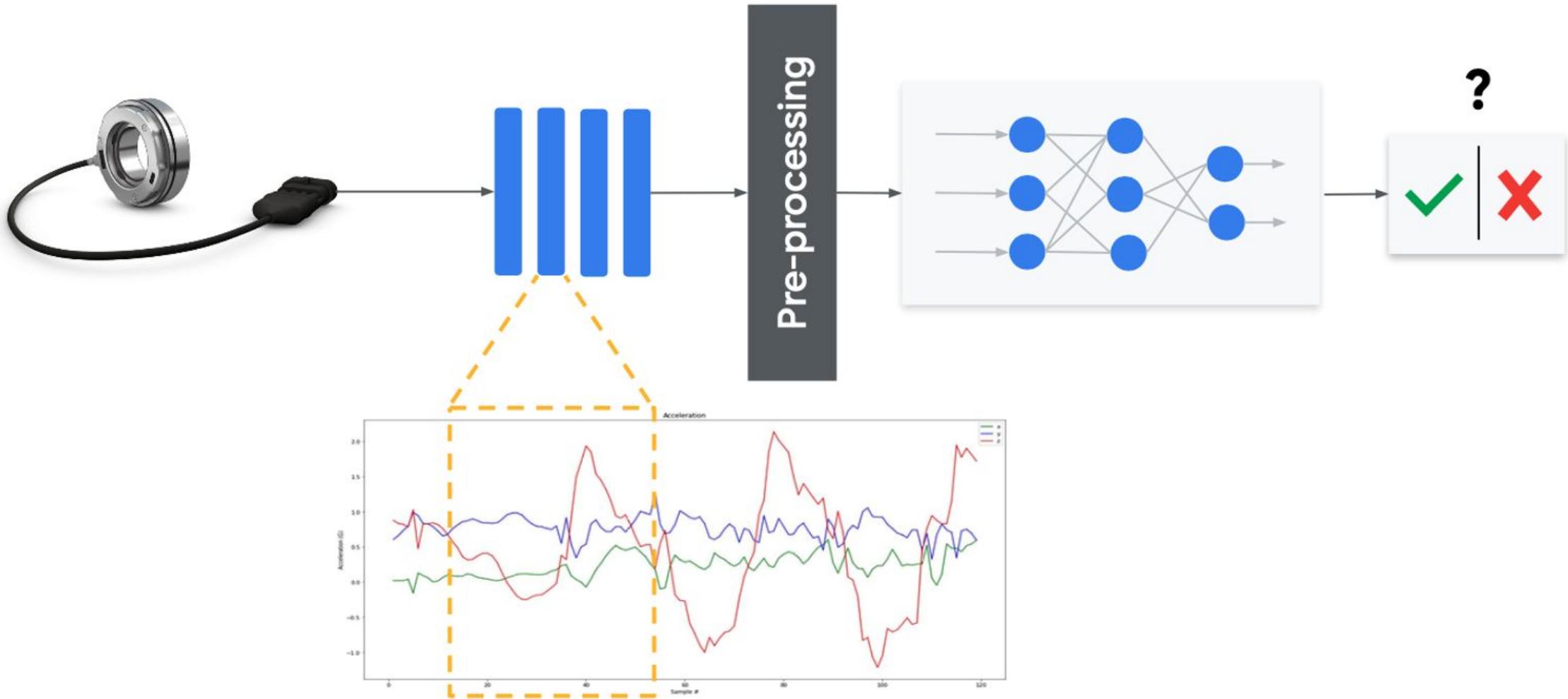


Ball Bearings



Accelerometer

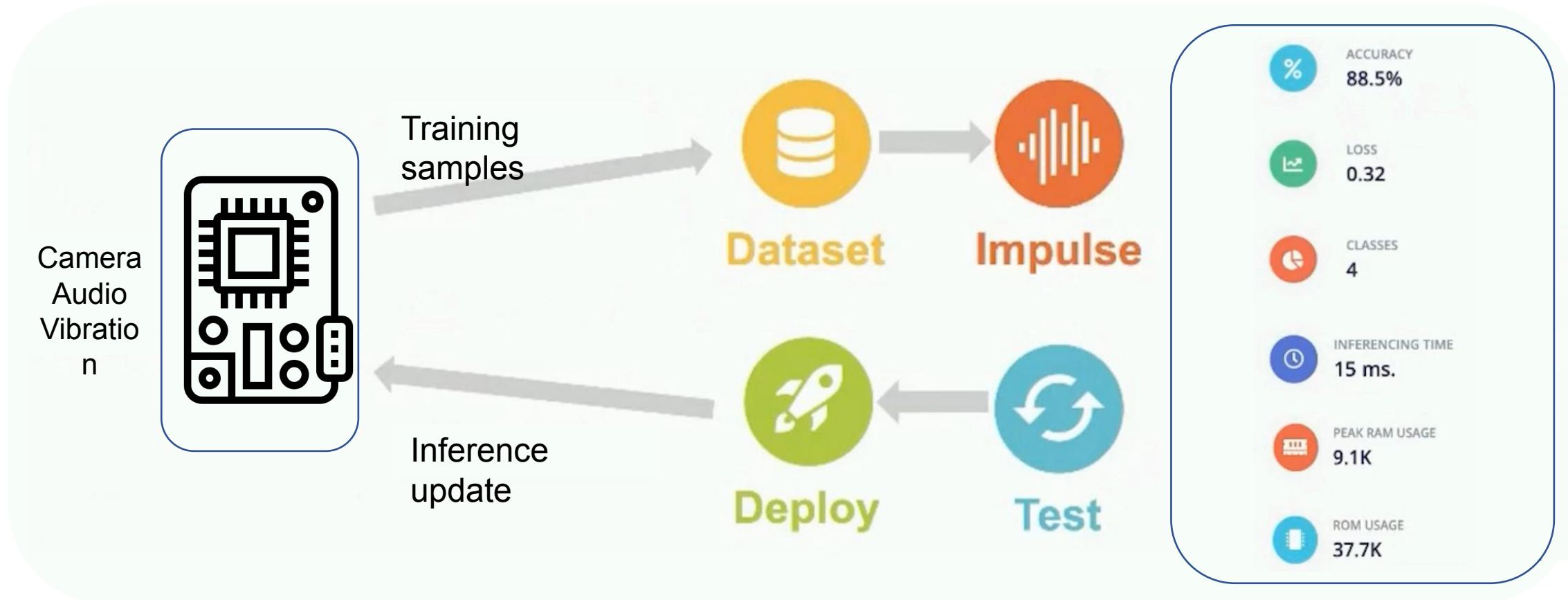
Anomaly Detection



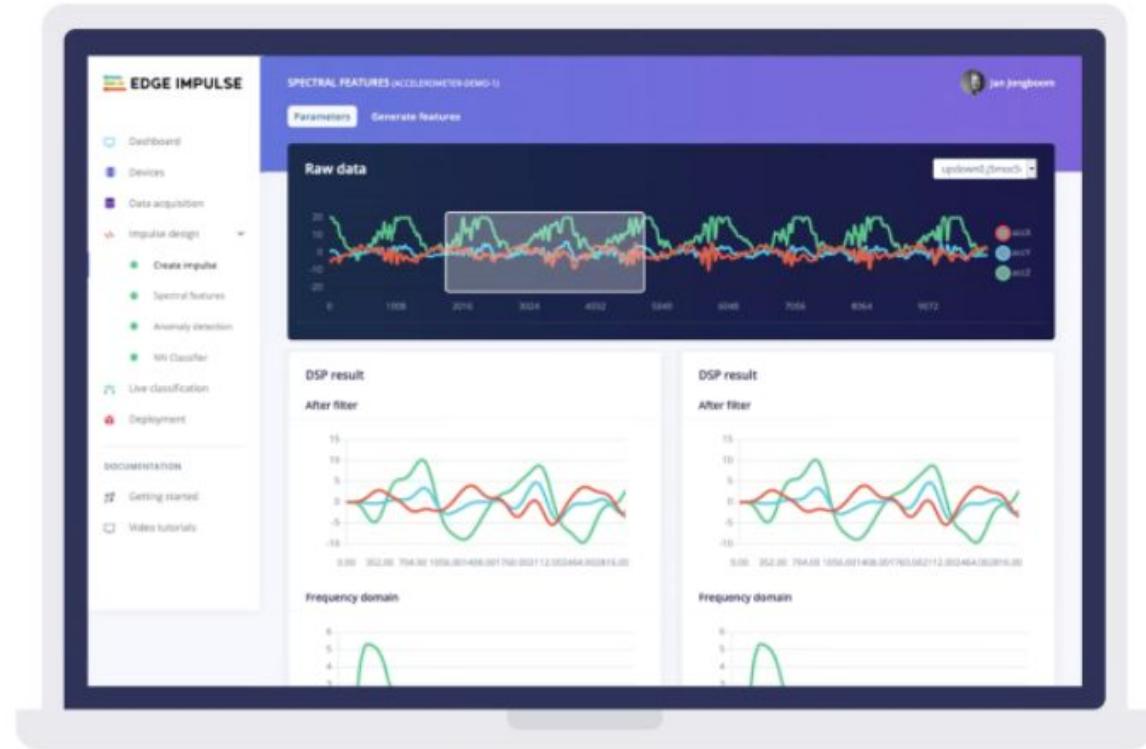
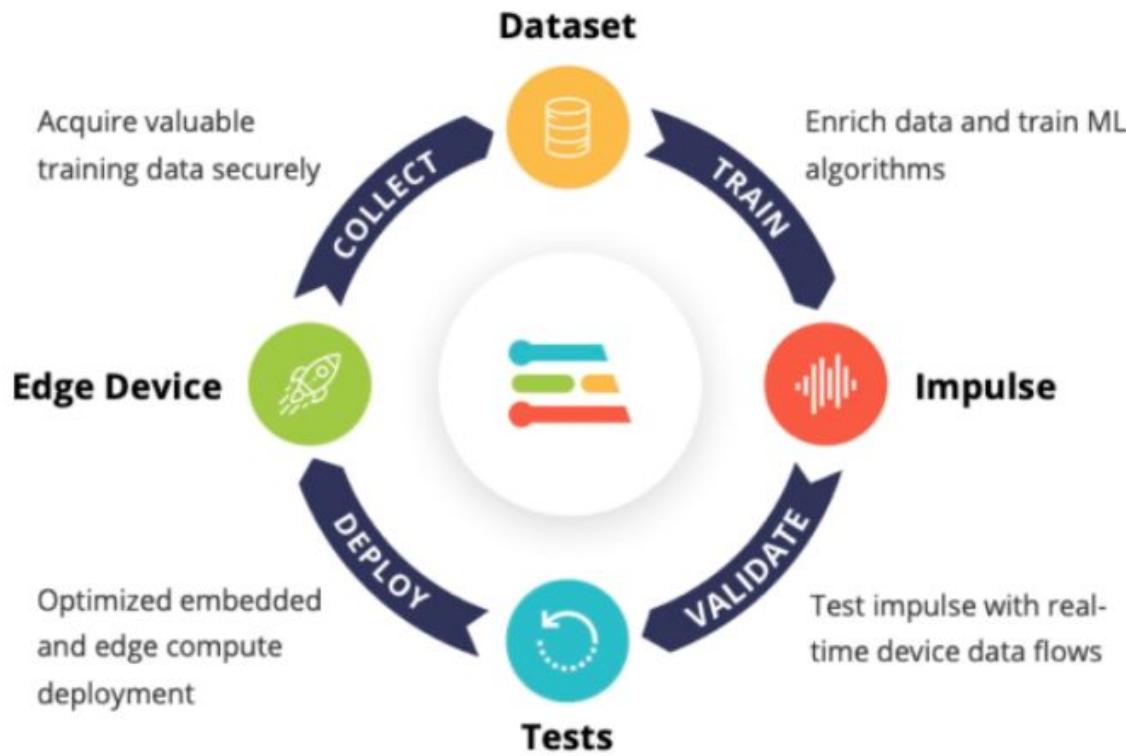
Data-driven engineering



Data-driven engineering



EI Studio - Embedded ML platform (“AutoML”)



Learn more at <http://edgeimpulse.com>



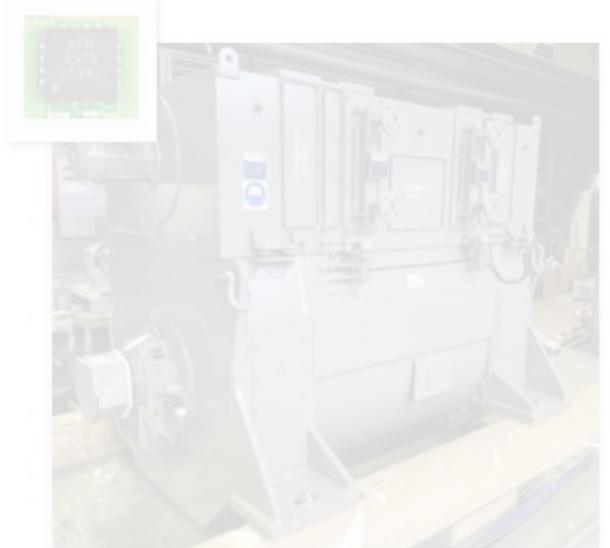
Sound

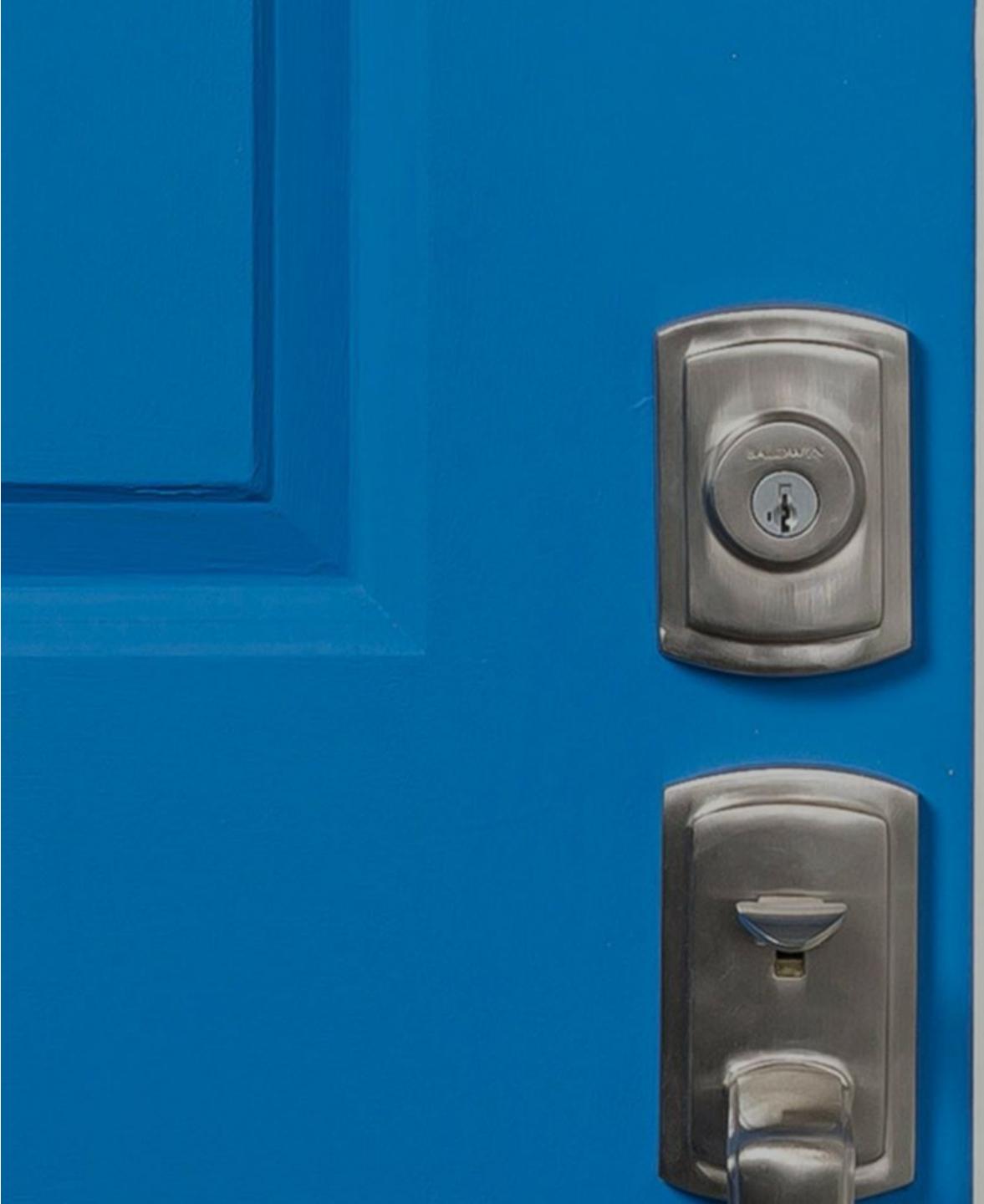


Vision

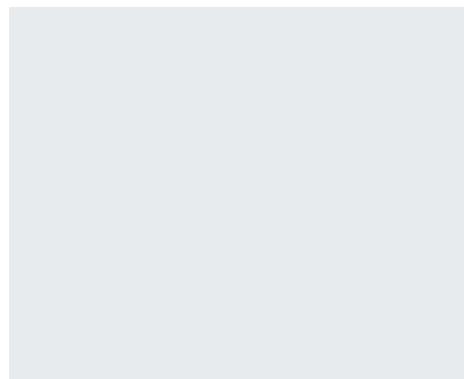


Vibration



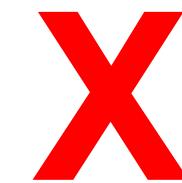
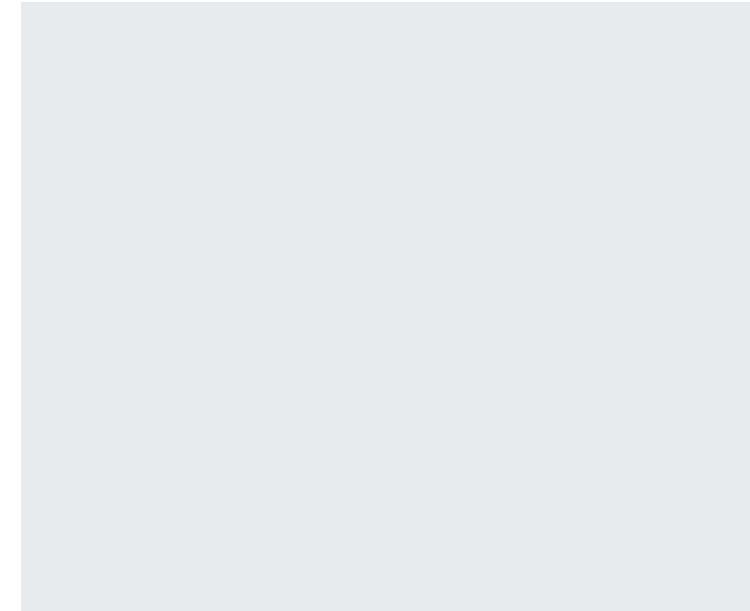


Person Detection

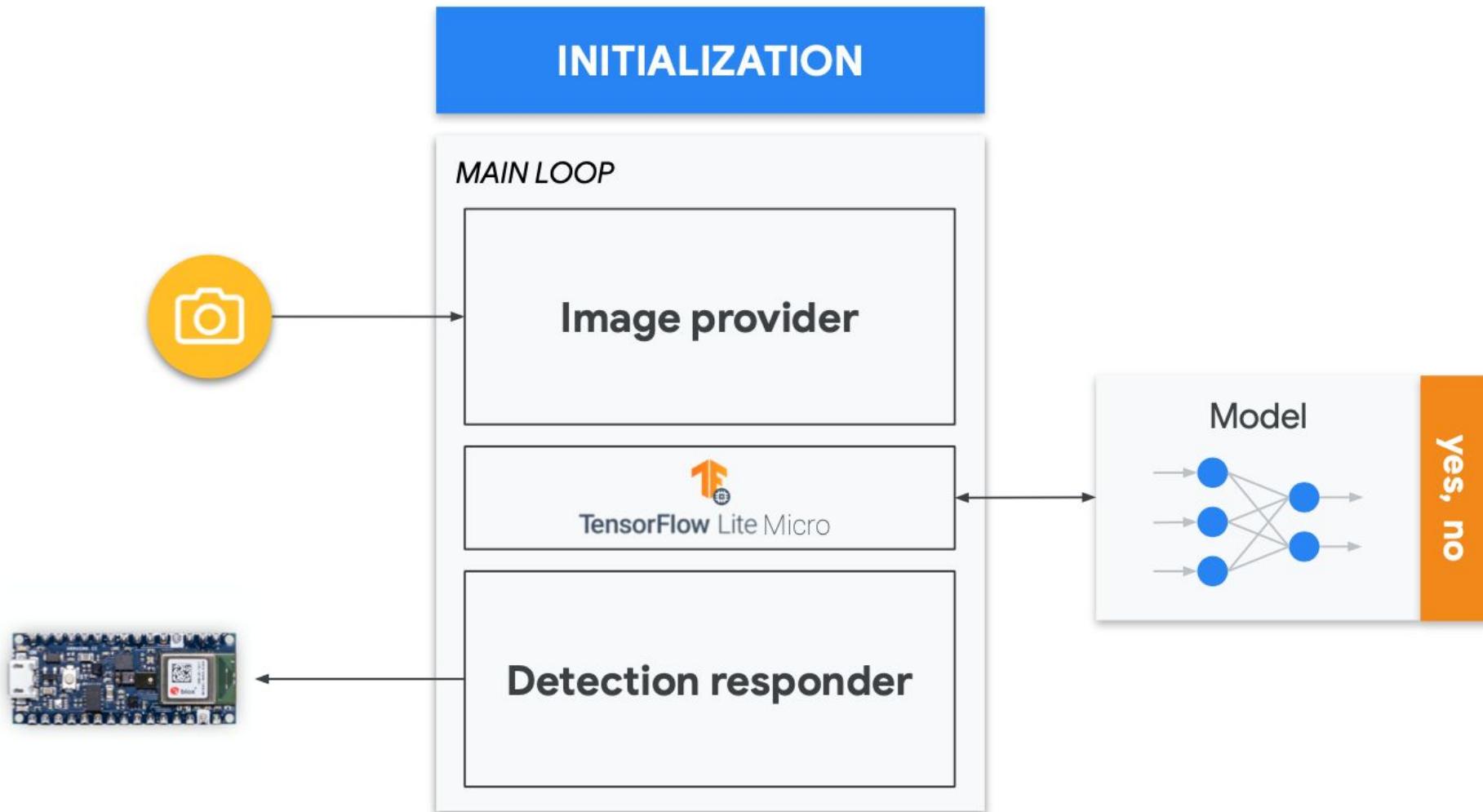


Mask Detection

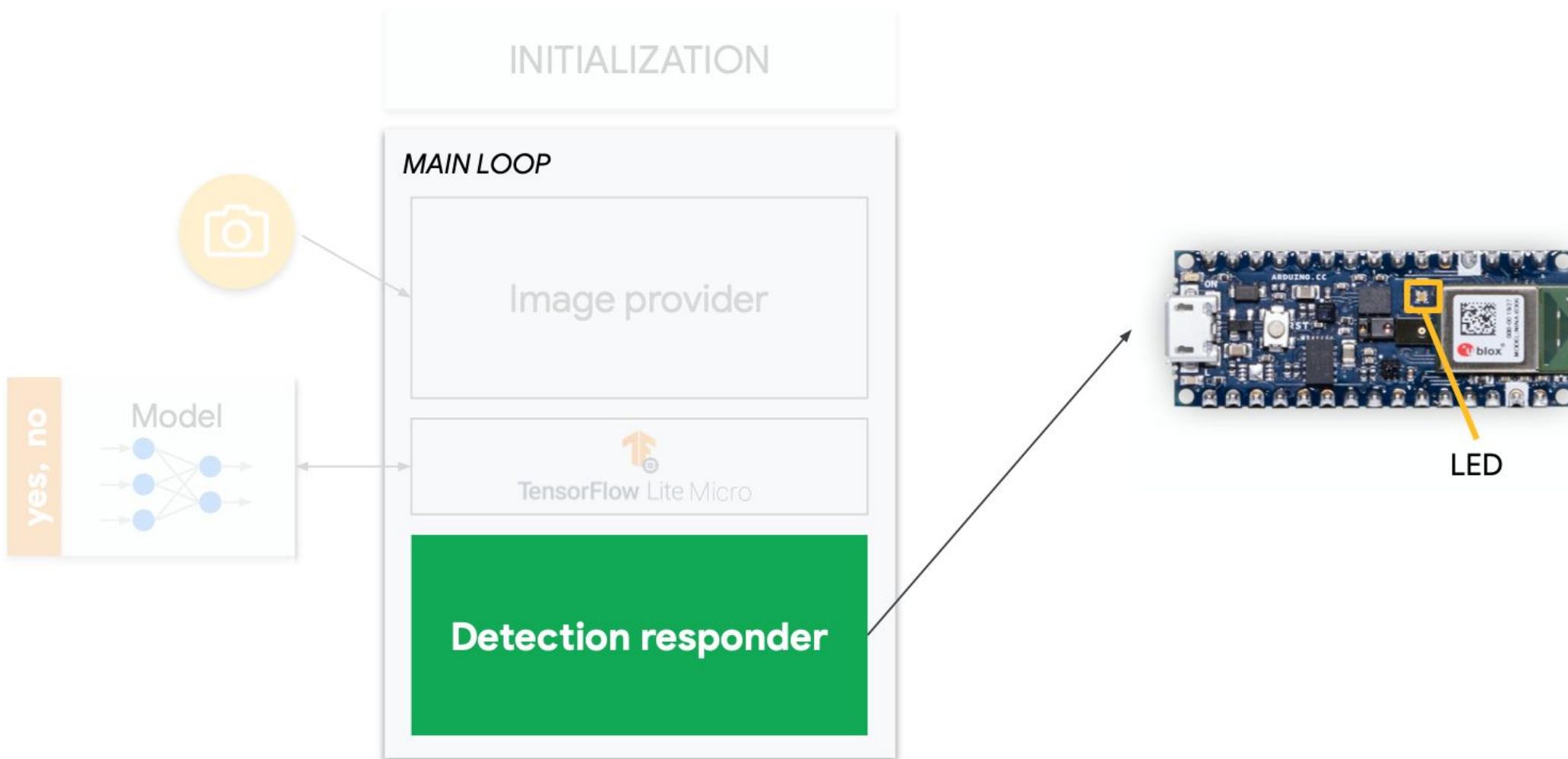




Person Detection Components



Post-processing



What AI really is?

What is (**Deep**) Machine Learning?

1. Machine Learning is a subfield of Artificial Intelligence focused on developing algorithms that learn to solve problems by analyzing data for patterns
2. **Deep Learning** is a type of Machine Learning that leverages **Neural Networks** and **Big Data**

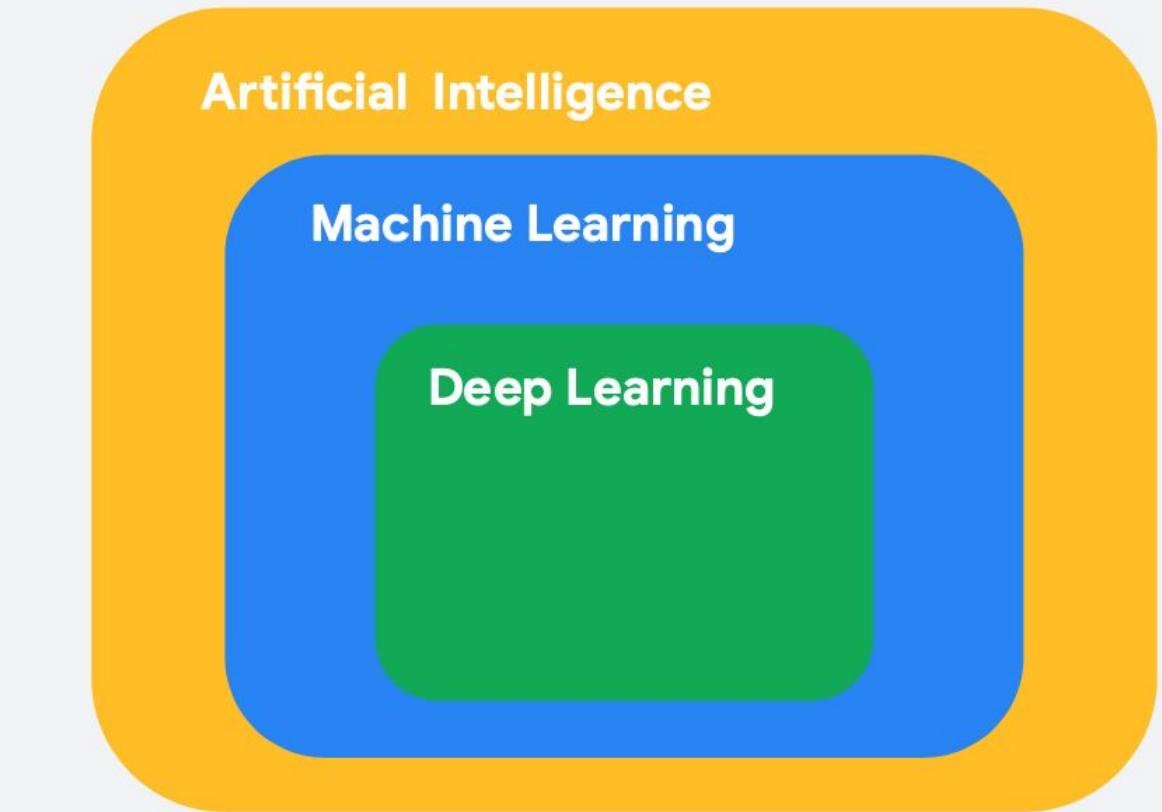
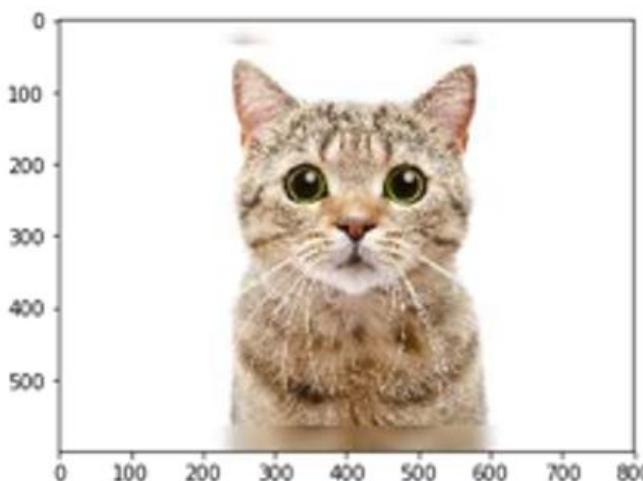
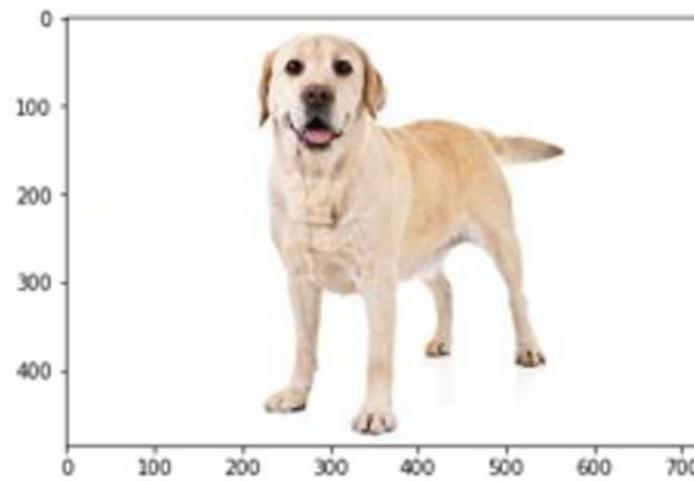


Image Classification

[PREDICTION]	[Prob]
Egyptian cat	: 64%
tabby	: 14%
bucket	: 3%



[PREDICTION]	[Prob]
Labrador retriever	: 83%
golden retriever	: 13%
bloodhound	: 0%



[PREDICTION]	[Prob]
German shepherd	: 60%
dhole	: 16%
malinois	: 7%



Object Detection

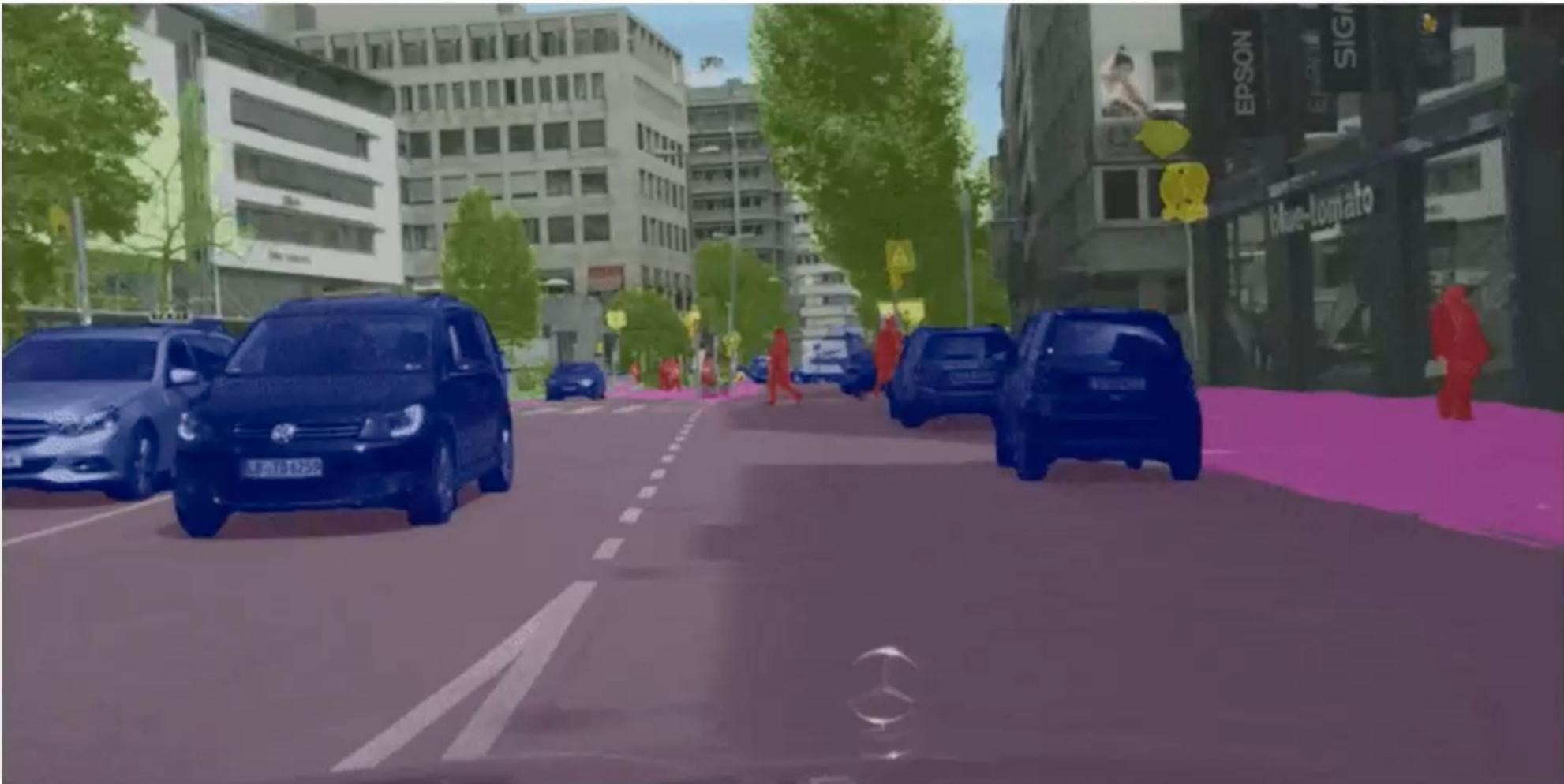


Photos



Live Video

Segmentation

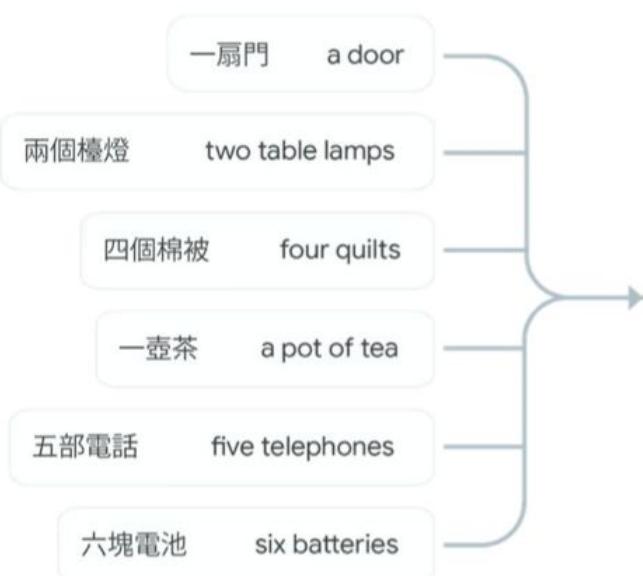


Pose Estimation



Machine Translation

1 Upload translated language pairs



2 Train your model



AutoML
Translation

3 Evaluate



Recommendations

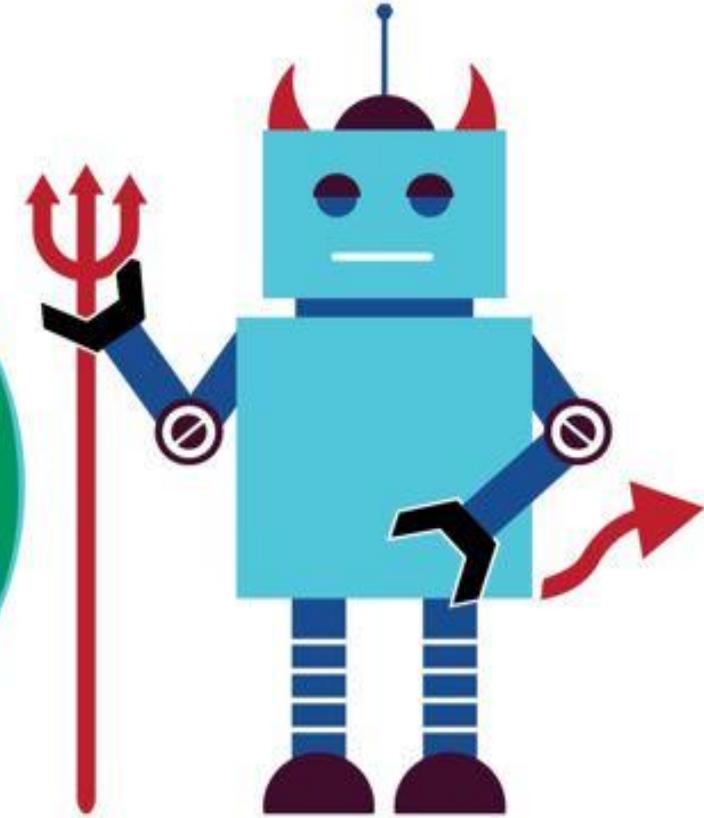
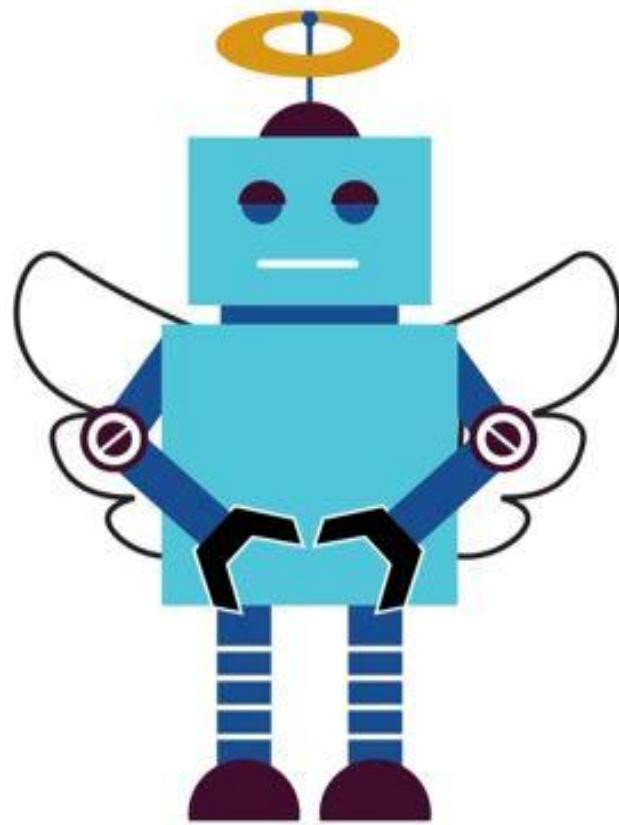
		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

General AI does not exist (yet)

Dedicated ML Applications

- Image Classification
- Object Detection
- Pose Estimation
- Voice Recognition
- Gesture Recognition
- Anomaly Detection
- Natural Language Processing (**NLP**)

Responsible AI



Reading Material

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 \(Coursera\)](#)
- [Text Book: "TinyML" by Pete Warden, Daniel Situnayake](#)

I want to thank Shawn Hymel and Edge Impulse, Pete Warden and Laurence Moroney from Google and specially Harvard professor Vijay Janapa Reddi, Ph.D. student Brian Plancher and their staff 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
And stay safe!

