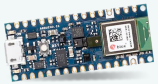
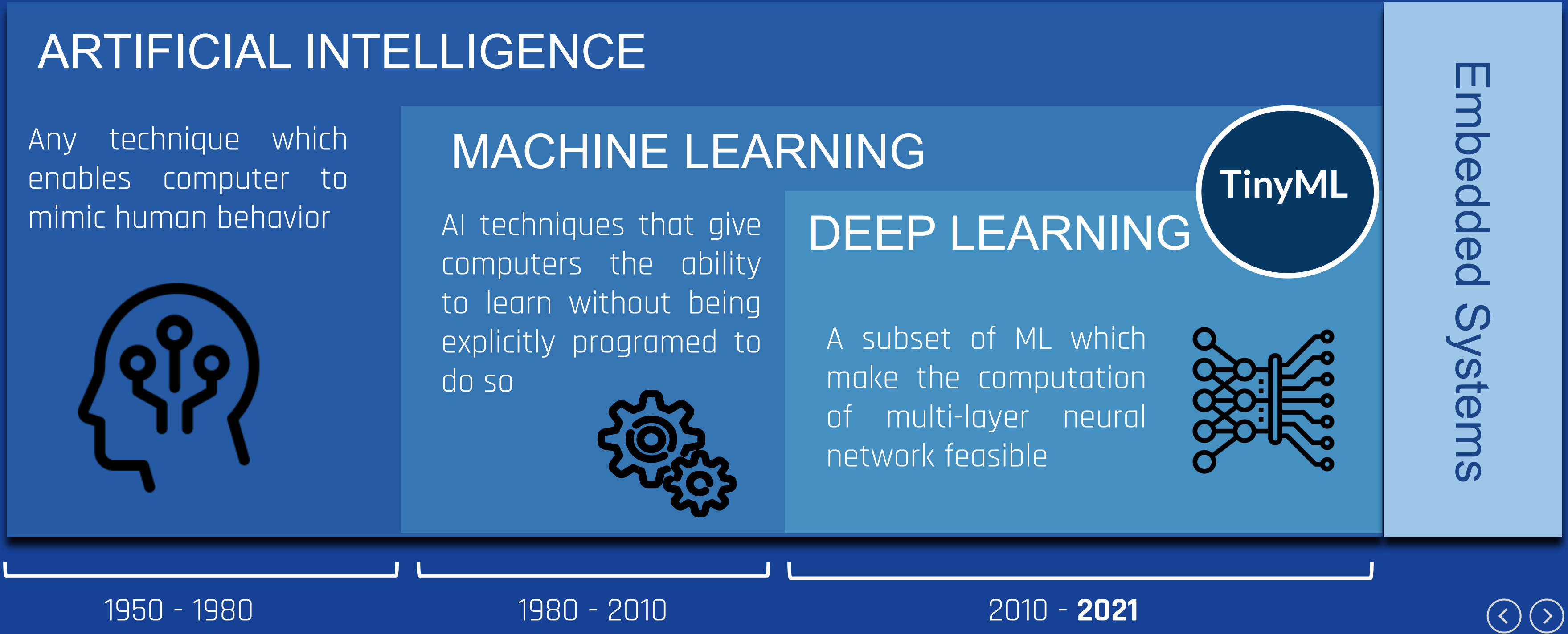
TinyML - the convergence between Machine Learning and the Internet of Things

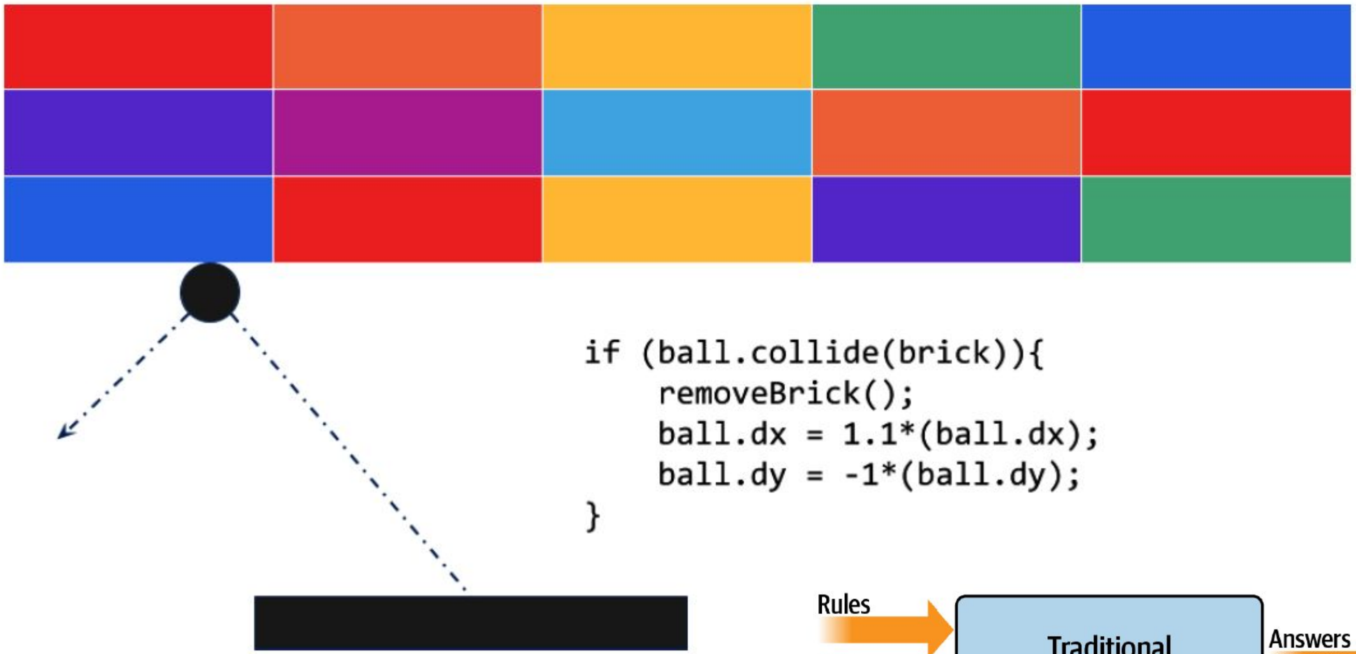
new opportunities are coming



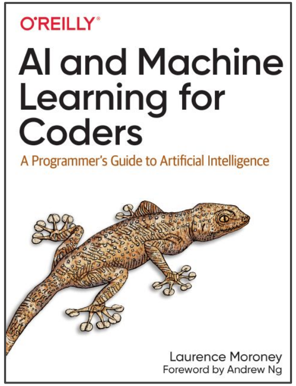
ARTIFICIAL INTELLIGENCE



What is Machine Learning?



```
if (ball.collide(brick)){
    removeBrick();
    ball.dx = 1.1*(ball.dx);
    ball.dy = -1*(ball.dy);
}
```



Limitations of traditional programming

<activity detection>

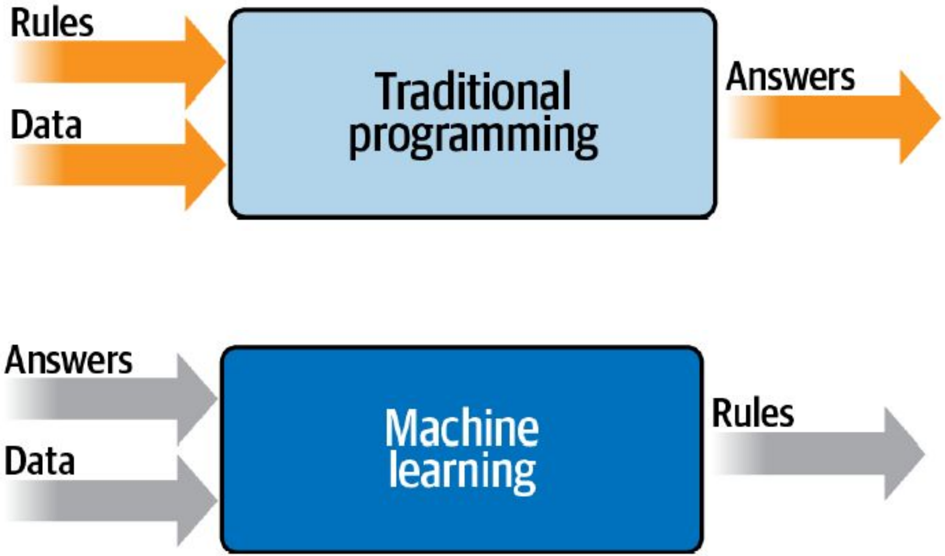
```
if(speed<4){
     status=WALKING;
}
```

```
if(speed<4){
     status=WALKING;
} else {
     status=RUNNING;
}
```

```
if(speed<4){
     status=WALKING;
} else if(speed<12){
     status=RUNNING;
} else {
     status=BIKING;
}
```

```
// ???
```

From programming to learning



From coding to ML

<gathering and label data>

01010010101001010101
10010101010010111101
01001010100101010010
01010010101001010101

Label = WALKING

10101001010010101010
01010100100100100010
00100111110101011111
10101001001111010111

Label = RUNNING

10010100111110101010
11010101110101011101
10101011110101010111
11111100011110101010

Label = BIKING

11111111110100111010
00111110101111101010
01011101010101011101
10101010101001111101

Label = GOLFING

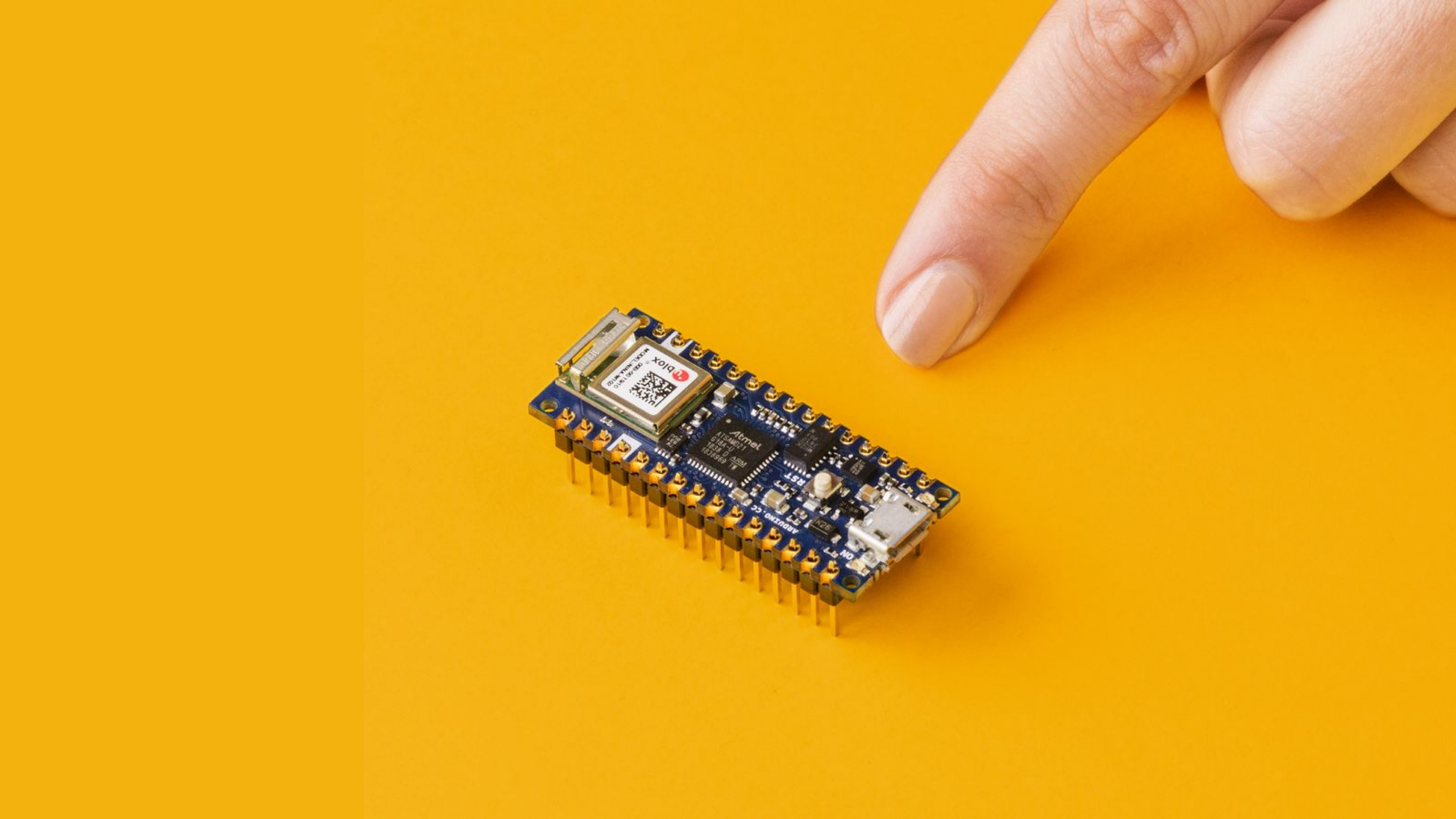
What is TinyML?

Tiny Machine Learning (TinyML) is a **fast-growing field of machine learning** technologies and applications including **algorithms, hardware and software** capable of performing **on-device sensor data analytics** (vision, audio, IMU, biomedical, etc.) at **extremely low power**, typically in the mW range and below, and hence enabling a variety of **always-on-use-cases** and targeting **battery-operated devices**.

OK Google

Alexa

Hey Siri


blox
Atmel
ARDUINO.CC
RST
ON

Smart OBD-II Scanners

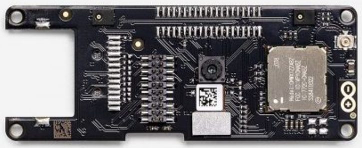
Pavement Identification

Heterogeneous Pavement

Outlier Detection

Pedro Andrade et al. An unsupervised TinyML approach applied for pavement anomalies detection under the IoIV. TBP, Metroind 2021

Portenta H7



Arduino Portenta Vision Shield - Lora

Preservation of wildlife

https://tago.io/blog/customer-nortronic/

Wildlife Monitoring

Risk Monitoring	Know when an animal is moving into a high-risk area and send real-time notifications to park rangers.
Conflict Monitoring	Sense and alert when a specie is heading into an area where farmers live.
Activity Monitoring	Classify the general behavior of the animal, such as when it is drinking, eating, sleeping, etc.
Communication Monitoring	Listen for vocal communication between animals via the onboard microphone

Google Trends

tinyml

Google Scholar

tinyml

Articles

About 306 results (0.03 sec)

My profile

Any time
Since 2021
Since 2020
Since 2017
Custom range...

Sort by relevance
Sort by date

include patents
include citations

Create alert

Benchmarking TinyML systems: Challenges and direction — [PDF] arxiv.org

CR Banbury, VJ Reddi, M Lam, W Fu, A Fazel… - arXiv preprint arXiv …, 2020 - arxiv.org

Recent advancements in ultra-low-power machine learning (**TinyML**) hardware promises to unlock an entirely new class of smart applications. However, continued progress is limited by the lack of a widely accepted benchmark for these systems. Benchmarking allows us to …

Cited by 30 Related articles All 3 versions

TinyML-Enabled Frugal Smart Objects: Challenges and Opportunities — [PDF] ieee.org

R Sanchez-Iborra, AF Skarmeta - IEEE Circuits and Systems …, 2020 - ieeexplore.ieee.org

The **TinyML** paradigm proposes to integrate Machine Learning (ML)-based mechanisms within small objects powered by Microcontroller Units (MCUs). This paves the way for the development of novel applications and services that do not need the omnipresent …

Cited by 16 Related articles

Tensorflow lite micro: Embedded machine learning on tinyml systems — [PDF] arxiv.org

R David, J Duke, A Jain, VJ Reddi, N Jeffries… - arXiv preprint arXiv …, 2020 - arxiv.org

Deep learning inference on embedded devices is a burgeoning field with myriad applications because tiny embedded devices are omnipresent. But we must overcome major challenges before we can benefit from this opportunity. Embedded processors are severely …

Cited by 21 Related articles All 2 versions

[PDF] **TinyML: Meta-data for Wireless Networks** — [PDF] berkeley.edu

N Ota, WTC Kramer - University of California at Berkeley …, 2003 - people.eecs.berkeley.edu

The growing number of deployed heterogeneous networks and applications is resulting in a sporadic isolated embedded sensor network environment. The **TinyML** project addresses the need for an embedded sensor network standardized “markup language” for …

Cited by 14 Related articles All 2 versions

IEEE.org | IEEE Xplore | IEEE-SA | IEEE Spectrum | More Sites   SUBSCRIBE   Cart   Create Account | Personal Sign In

IEEE Xplore®   Browse   My Settings   Help   Institutional Sign In   IEEE

All

ADVANCED SEARCH

Search within results

Per Page: 25 | Export | Set Search Alerts | Search History

Showing **1-11** of **11** for **tinyml**

Conferences (8)   Journals (2)   Magazines (1)

Show

All Results

Open Access Only

Year

Single Year | Range

2020   2021

Select All on Page

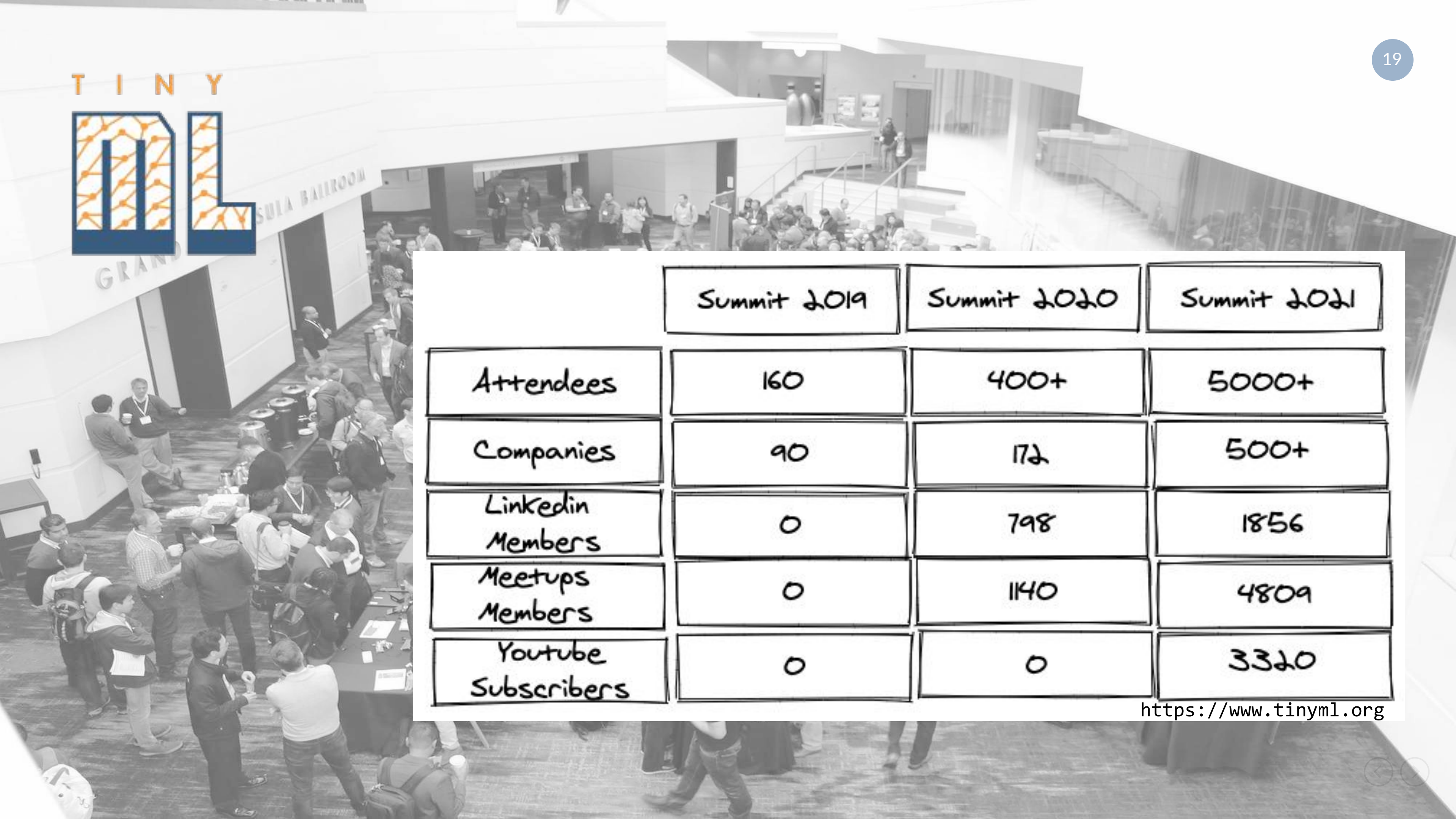
Sort By: Relevance

TinyML-Enabled Frugal Smart Objects: Challenges and Opportunities
Ramon Sanchez-Iborra; Antonio F. Skarmeta
IEEE Circuits and Systems Magazine
Year: 2020 | Volume: 20, Issue: 3 | Magazine Article | Publisher: IEEE
Cited by: Papers (5)

Abstract   html   (1498 Kb)

Adaptive Traffic Control With TinyML
A. Navaas Roshan; B. Gokulapriyan; C. Siddarth; Priyanka Kokil
2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)

Feedback



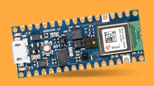
	Summit 2019	Summit 2020	Summit 2021
Attendees	160	400+	5000+
Companies	90	172	500+
Linkedin Members	0	798	1856
Meetups Members	0	1140	4809
Youtube Subscribers	0	0	3320

https://www.tinyml.org

Agenda Recap

Fundamentals of TinyML

Applications of TinyML

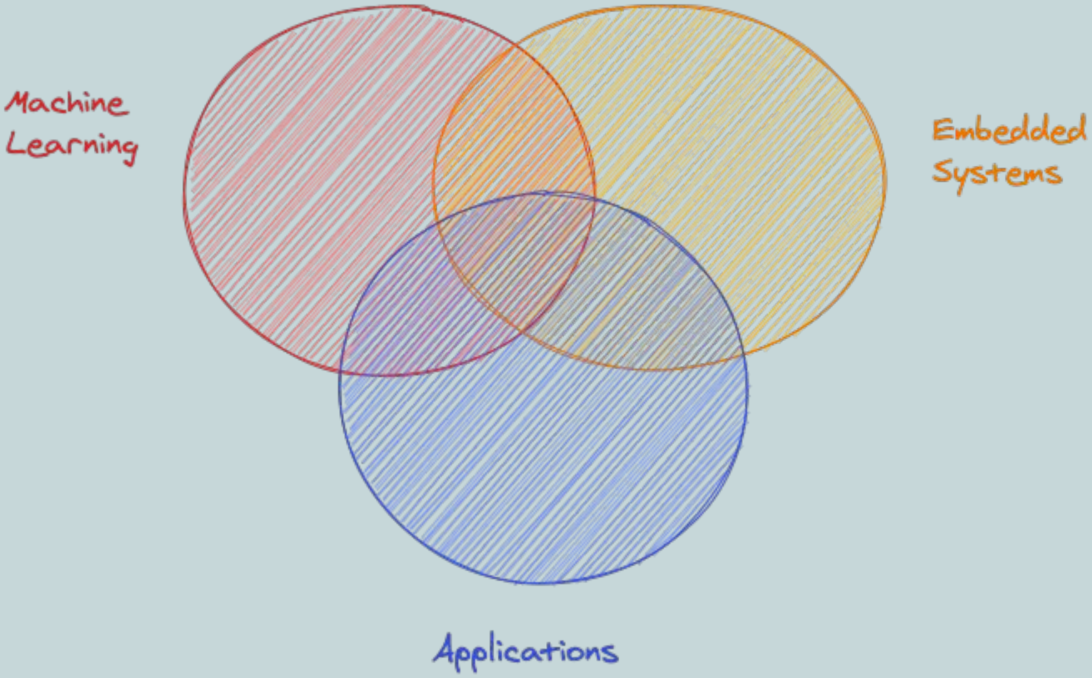


Deploying TinyML

Papers you should read:

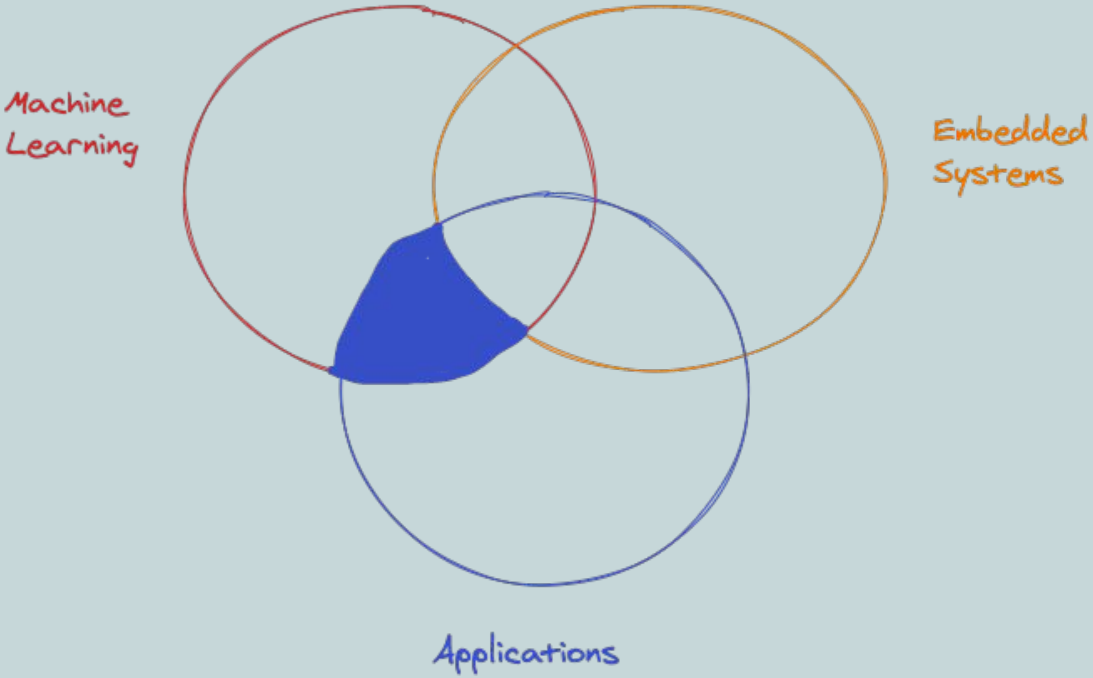
1. **MicroNets: Neural Network Architectures for Deploying TinyML Applications on Commodity Microcontrollers**
   a. https://arxiv.org/abs/2010.11267
2. **TensorFlow Lite Micro: Embedded Machine Learning on TinyML Systems**
   a. https://arxiv.org/abs/2010.08678
3. **Benchmarking TinyML Systems: Challenges and Direction**
   a. https://arxiv.org/abs/2003.04821
4. **Quantized Neural Network Inference with Precision Batching**
   a. https://arxiv.org/abs/2003.00822

TinyML Fundamentals



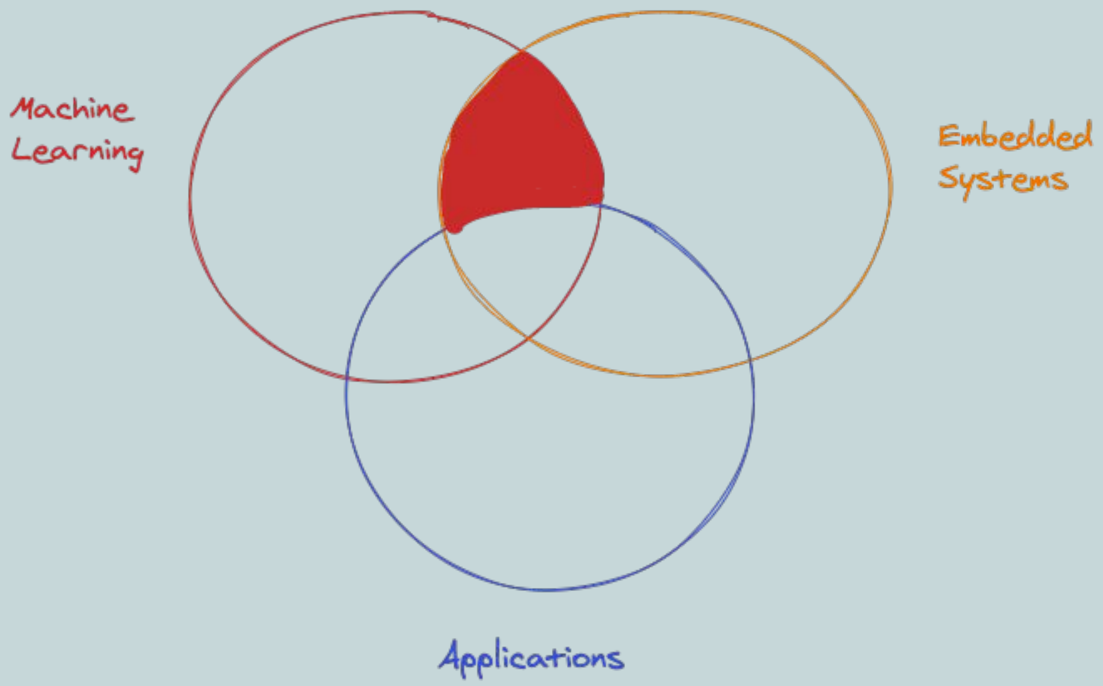
Interactions

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



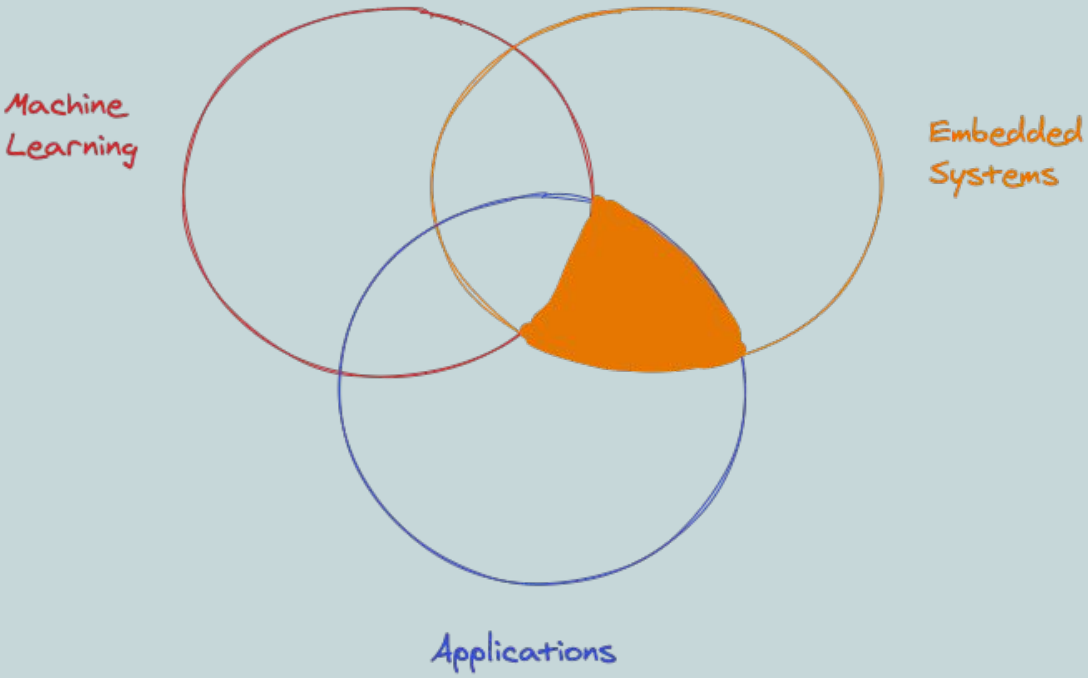
Interactions

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



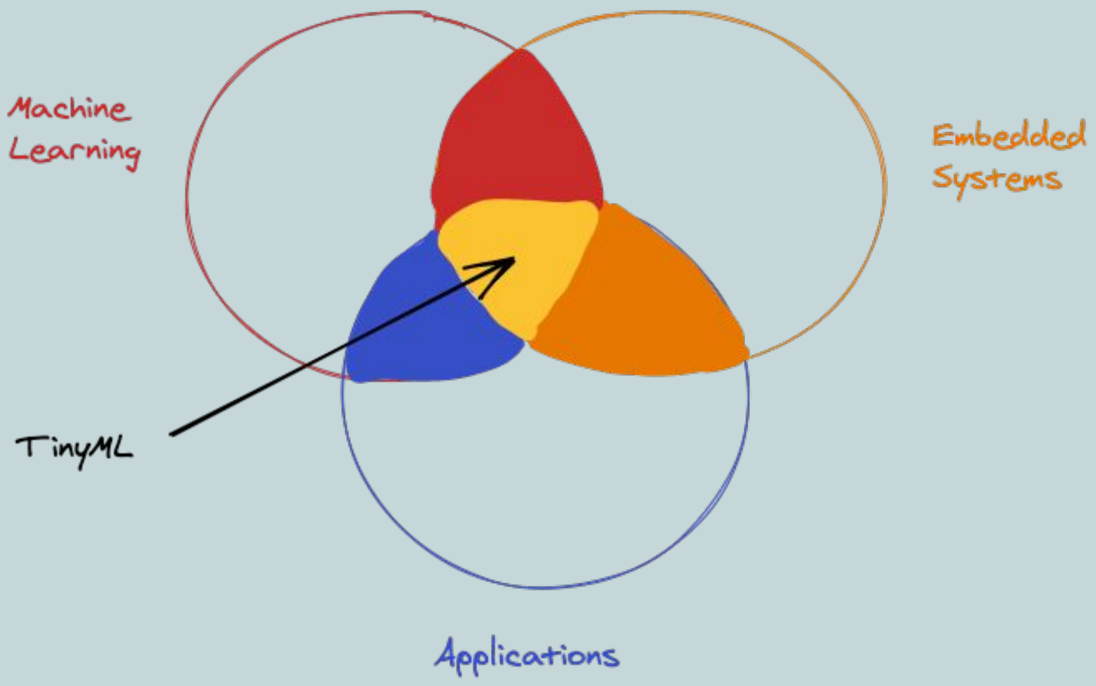
Interactions

What are the challenges with enabling ML on **Tiny**, resource-constrained embedded devices?



Interactions

What type of new use cases can we possibly enable on embedded system what 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**

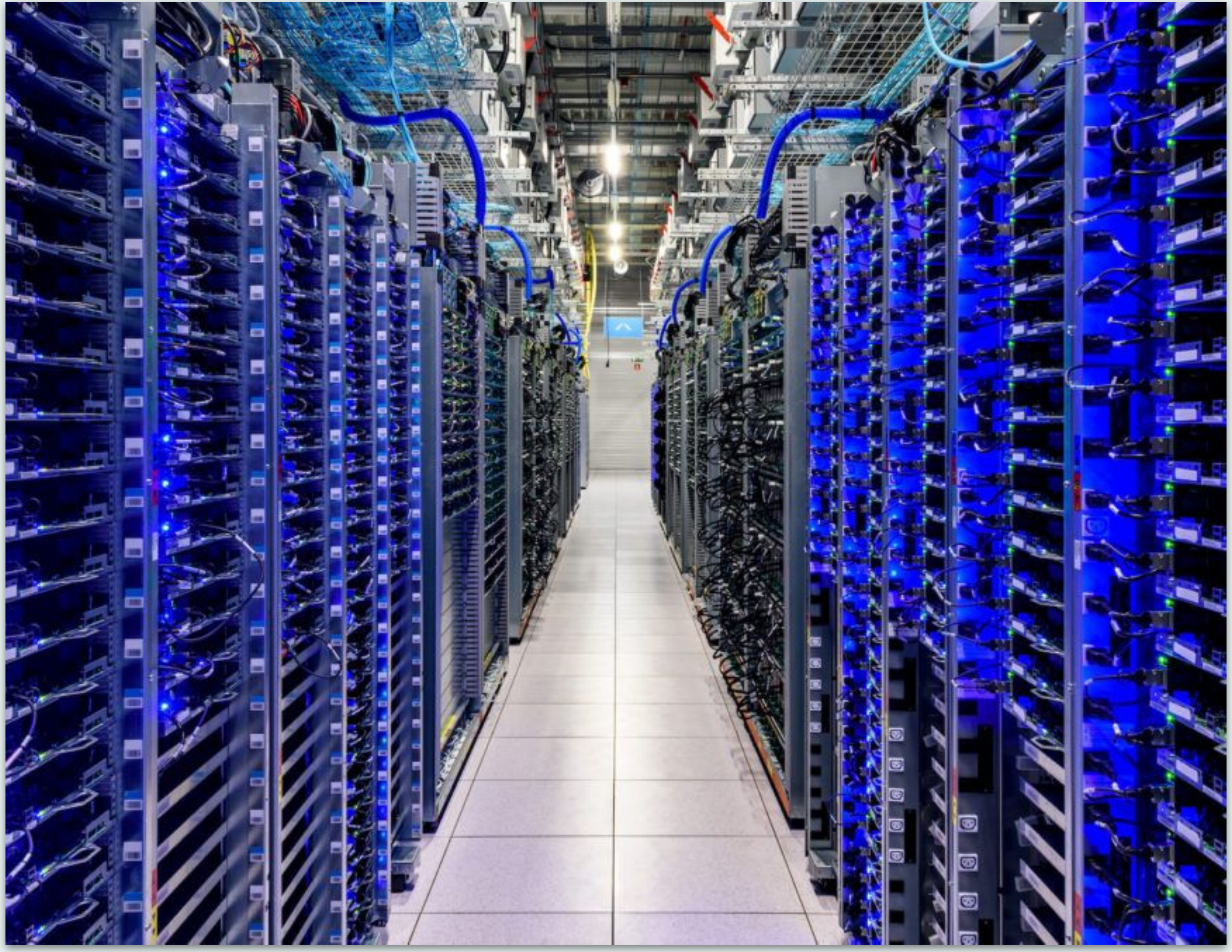
How am I going to get there?


ARDUINO
TINY MACHINE
LEARNING KIT
edX
TensorFlow
ARDUINO
NANO 33 SENSE
BOARD

TensorFlow

Why the future of machine learning is **tiny** and **bright**?

Machine Learning is here to stay



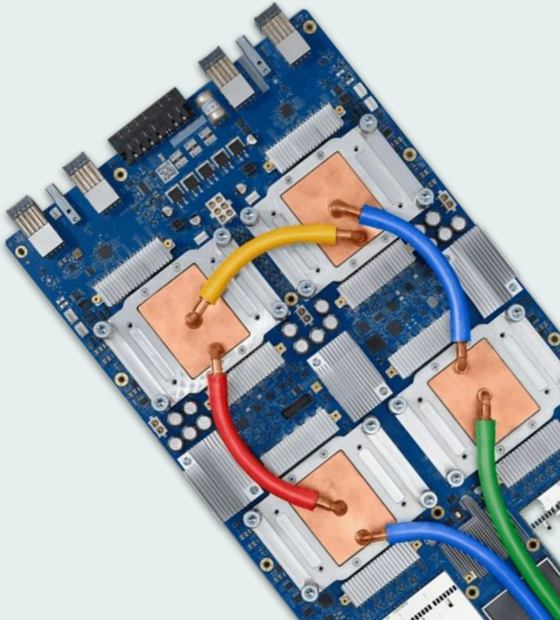
These capabilities require a remarkable of computing capabilities.

DataCenter

Graph Processing Units (GPU) vs Tensor Processing Units (TPU)

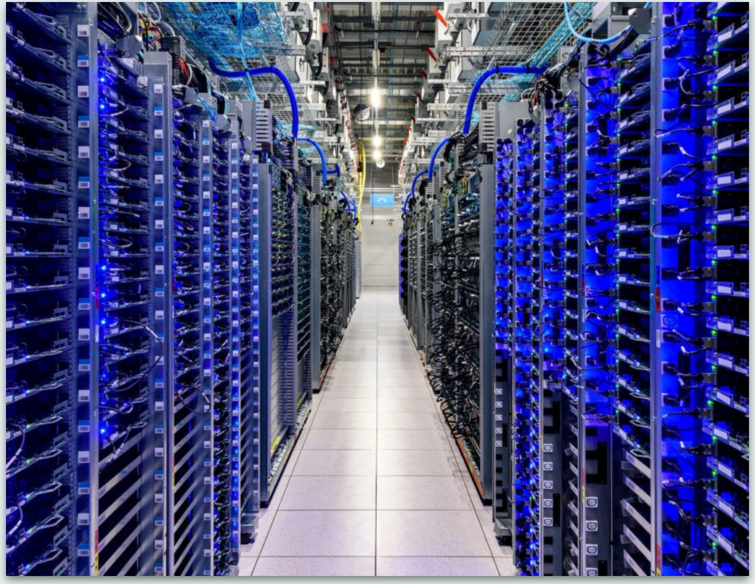
NVIDIA

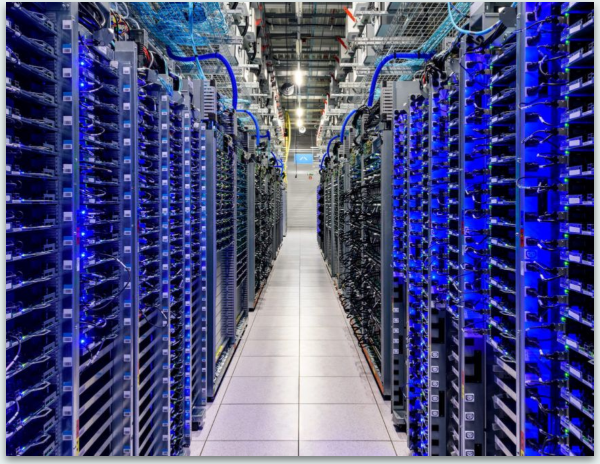
Google



Bigger is Not Always Better!!!

Interactivity? Responsiveness?

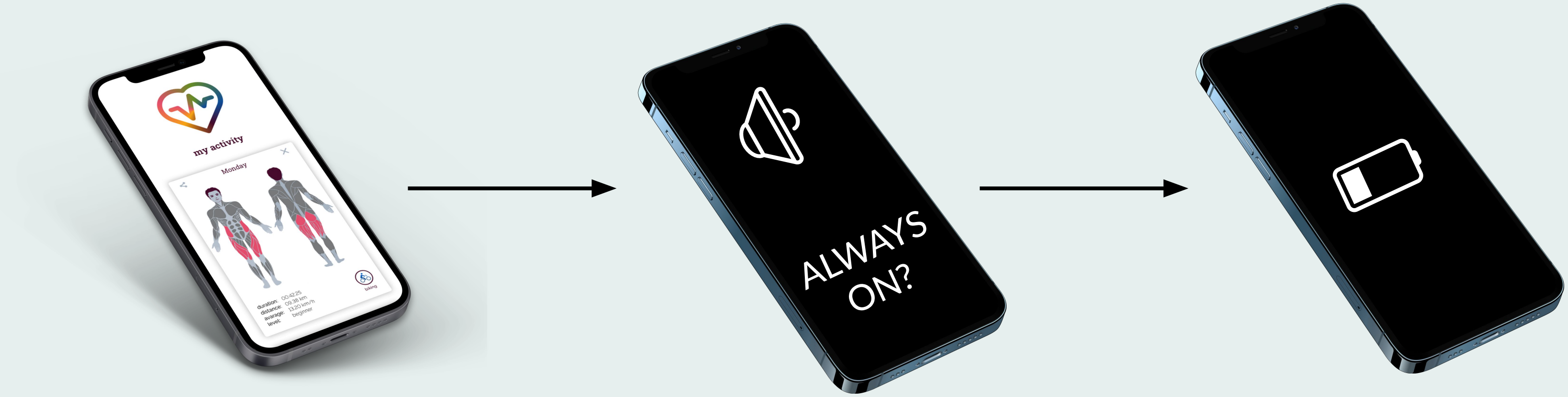




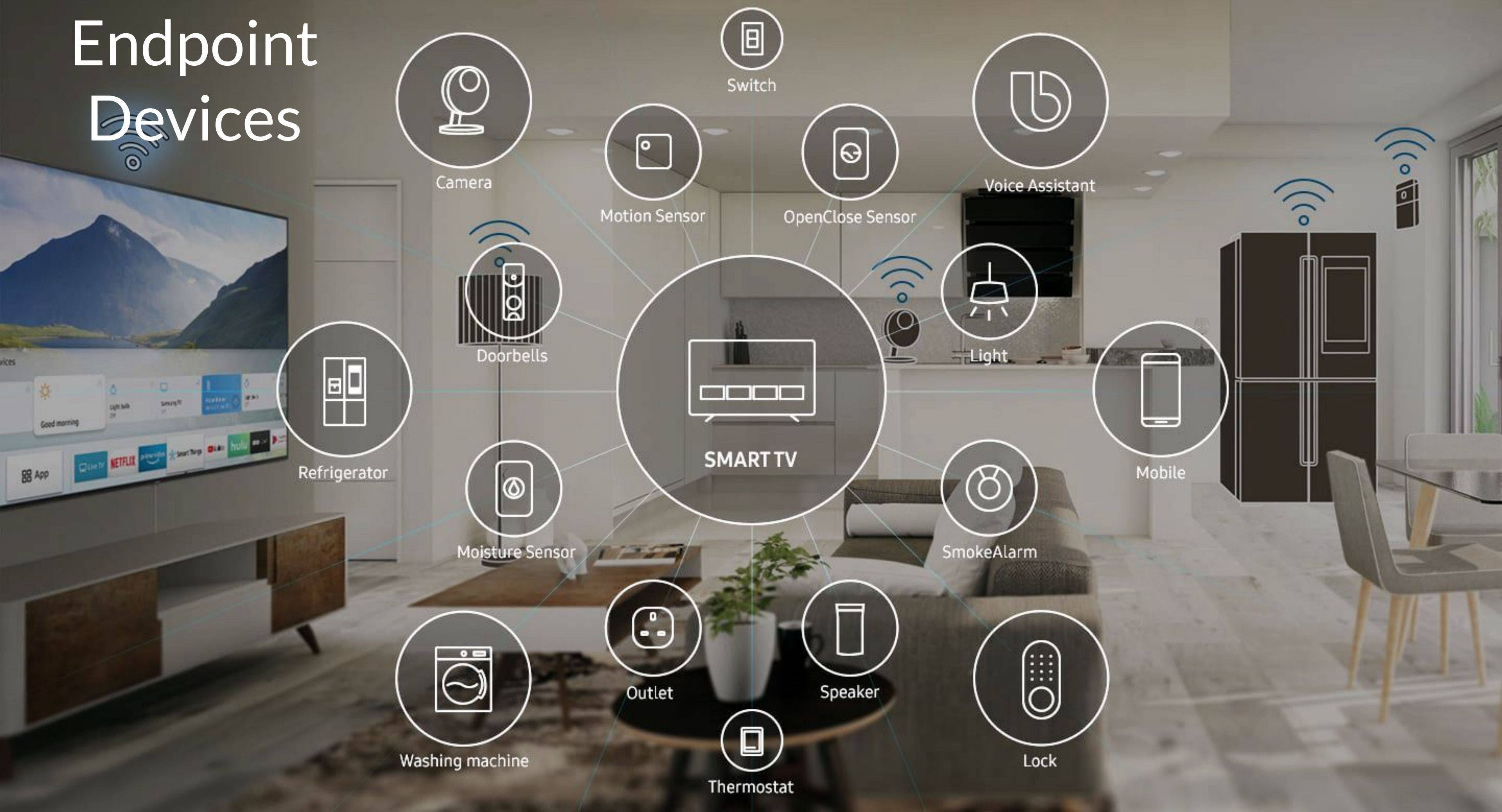
High Power
High Bandwidth
High Latency

Low Power
Low Bandwidth
Low Latency

Machine Learning running on phone all the time, always on?

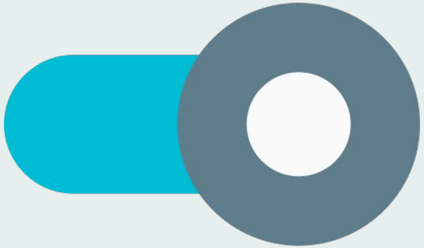


Endpoint Devices



A brief recap

- ML has several diverse applications in the real-world
- ML is increasingly moving from the cloud to endpoint devices
- Endpoint devices are everywhere around us



How do we enable TinyML?

What Makes TinyML?

Embedded Systems

Machine Learning

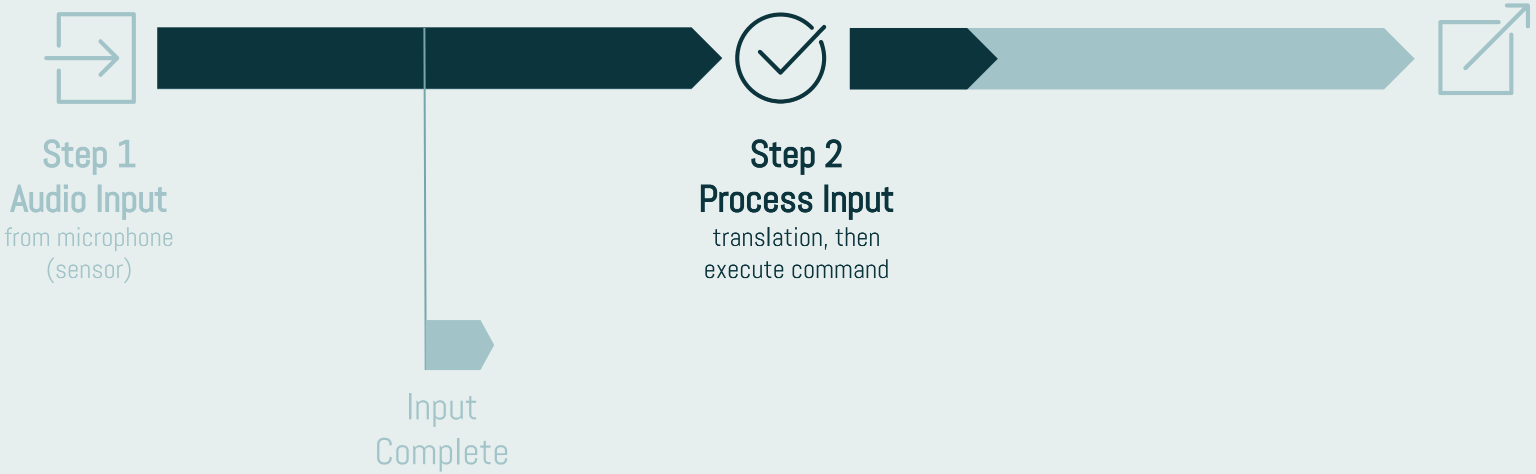
Okay,Google!

Hi, how can I help?

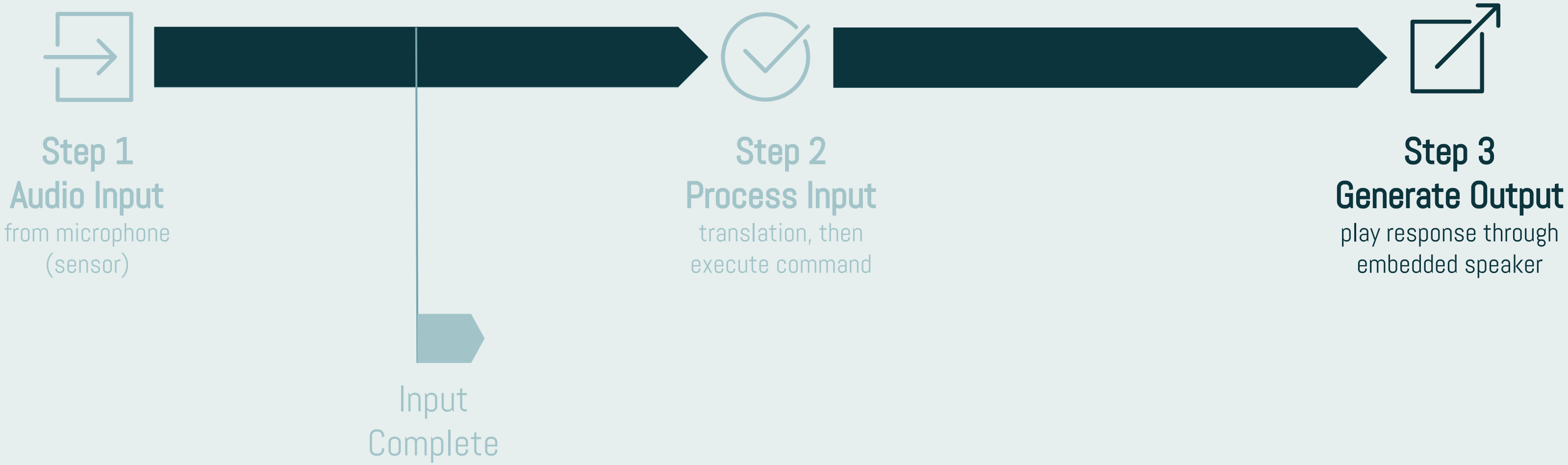
The Three Basic Steps

The Three Basic Steps

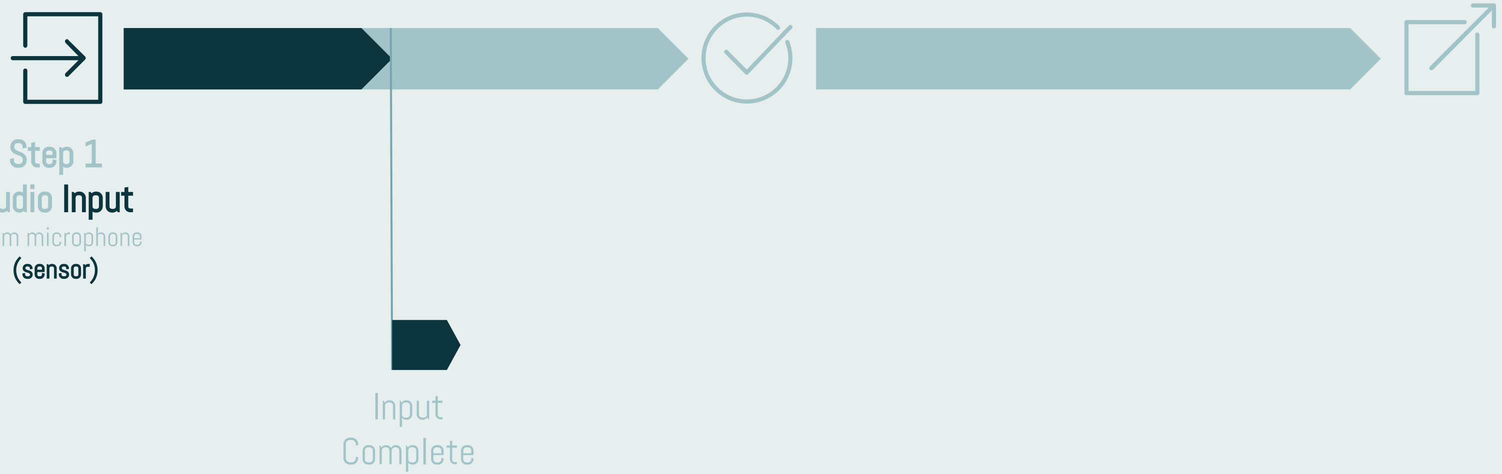
The Three Basic Steps



The Three Basic Steps



Input



Endpoints Have Sensors, Tons of Sensors

Motion Sensors,
Gyroscope, Magnetometer,
Radar, Accelerator

Acoustic Sensors,
Ultrasonic, Microphones,
Geophones, Vibrometers

Environment Sensors,
Temperature, Humidity
Pressure, IR, etc

Touchscreen Sensors,
Capacitive, IR

Image Sensors,
Thermal, Image

Biometric Sensors,
Fingerprint, Heart Rate

Force Sensors,
Pressure, Strain

Rotation Sensors,
Encoders

Endpoints Have Sensors, Tons of Sensors

Motion Sensors, Gyroscope, Magnetometer, Radar, Accelerator

Acoustic Sensors, Ultrasonic, Microphones, Geophones, Vibrometers

Environment Sensors, Temperature, Humidity Pressure, IR, etc

Touchscreen Sensors, Capacitive, IR

Image Sensors, Thermal, Image

Biometric Sensors, Fingerprint, Heart Rate

Force Sensors, Pressure, Strain

Rotation Sensors, Encoders

Biometric Sensors

Non-invasive Glucose Monitoring

Fingerprint + Photoplethysmography (PPG)

Endpoints Have Sensors, Tons of Sensors

Motion Sensors, Gyroscope, Magnetometer, Radar, Accelerator

Acoustic Sensors, Ultrasonic, Microphones, Geophones, Vibrometers

Environment Sensors, Temperature, Humidity Pressure, IR, etc

Touchscreen Sensors, Capacitive, IR

Image Sensors, Thermal, Image

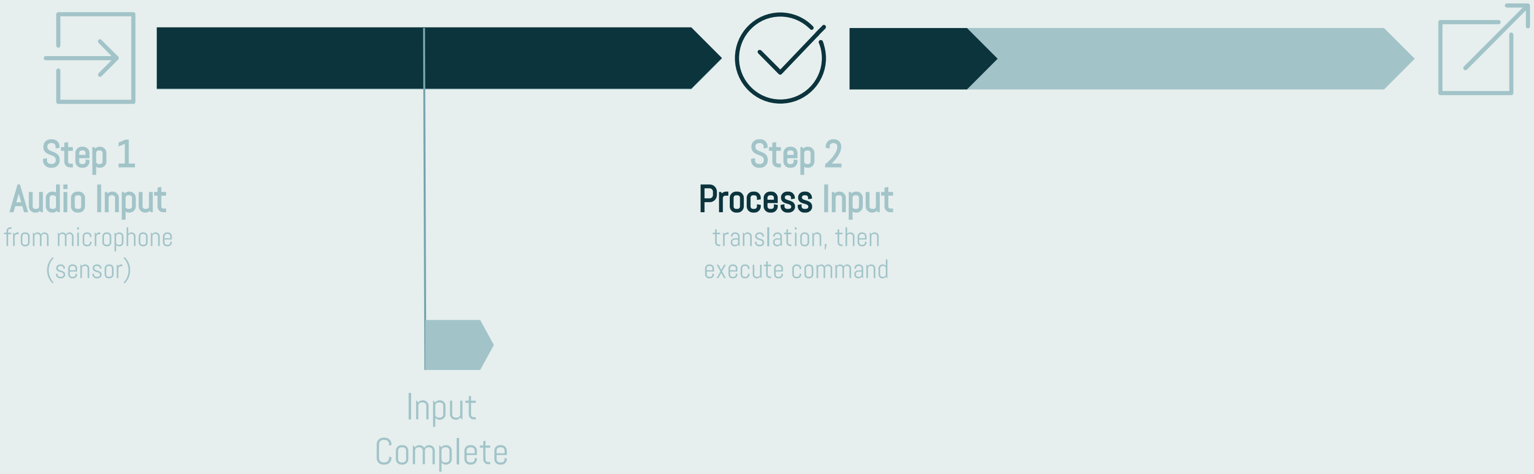
Biometric Sensors, Fingerprint, Heart Rate

Force Sensors, Pressure, Strain

Rotation Sensors, Encoders


ARDUINO
TINY MACHINE
LEARNING KIT
TensorFlow
NANO 33 SENSE
BOARD BLE WITH HEADERS

Processing


Step 1
Audio Input
from microphone
(sensor)
Input
Complete
Step 2
Process Input
translation, then
execute command

Thinking Big

BIG

GPU/CPU

561 mm^2

Thinking Small

BIG

GPU/CPU

561 mm^2

Mobile SoC

83mm^2

Thinking Tiny

BIG
GPU/CPU
561 mm^2

SMALL

Mobile SoC
83mm^2

Tiny

Apple 0778
30mm^2

Thinking Record-Breaking

BIG

GPU/CPU

561 mm^2

SMALL

Mobile SoC

83mm^2

Tiny

Apple 0778

30mm^2

World's smallest ARM-Powered MCU

48MHz, 32kB flash, 20-pin

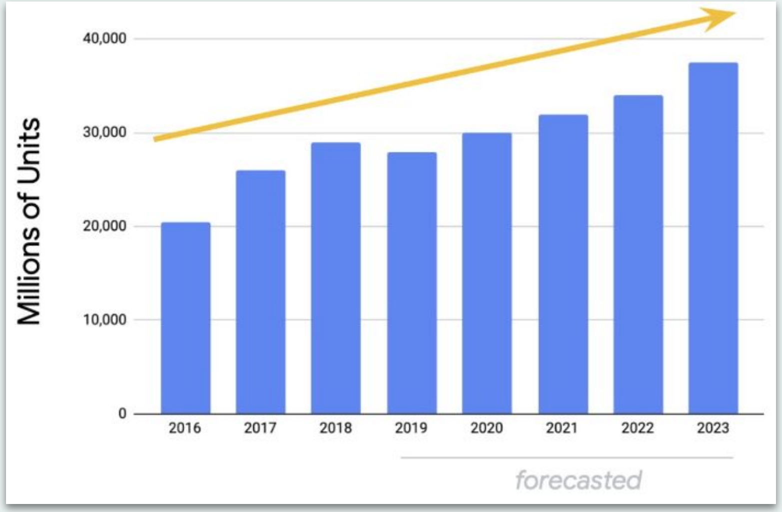
Kinetis KL03

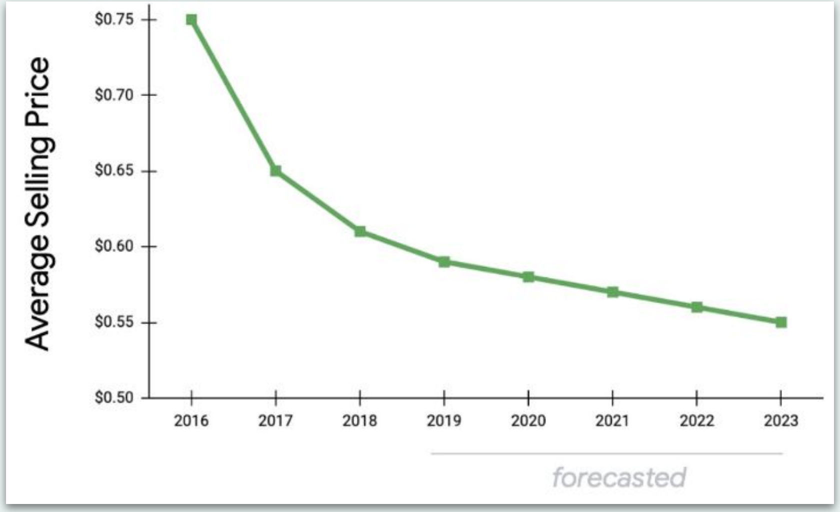
3.2mm^2


2.0 mm
1.6 mm

250 Billion MCU Today

MCU Demand and Pricing Forecast





Source: IC Insights

Comparing Power

BIG

GPU/CPU

561 mm^2

SMALL

Neural Decision Processor

Always-on deep learning

speech/audio recognition

Ultra low power, 128KB SRAM

12-pin, 2.52mm^2

300W

NVIDIA Tesla K80

3.64W

Apple A12

140 μW

Syntian NDP100

Comparing Power

Use case: button cell battery

Neural Decision Processor
Always-on deep learning
speech/audio recognition
Ultra low power, 128KB SRAM
12-pin, 2.52mm^2

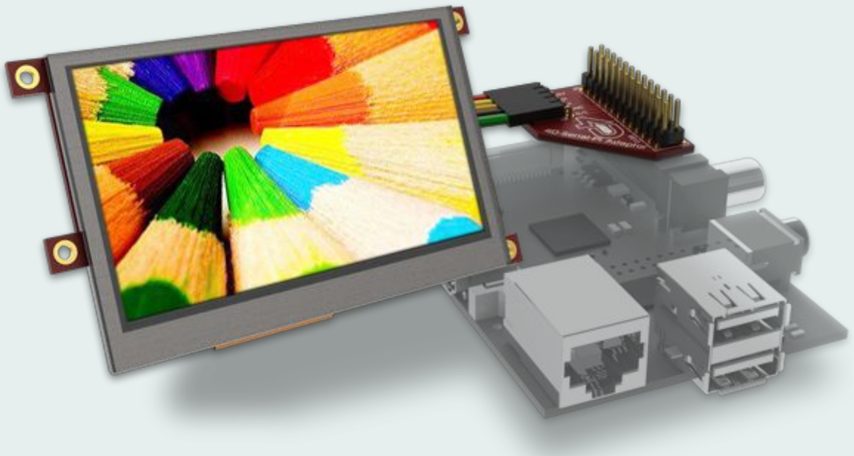
140 μW
Syntian NDP100

Output


Step 1
Audio Input
from microphone
(sensor)
Step 2
Process Input
translation, then
execute command
Step 3
Generate Output
play response through
embedded speaker
Input
Complete

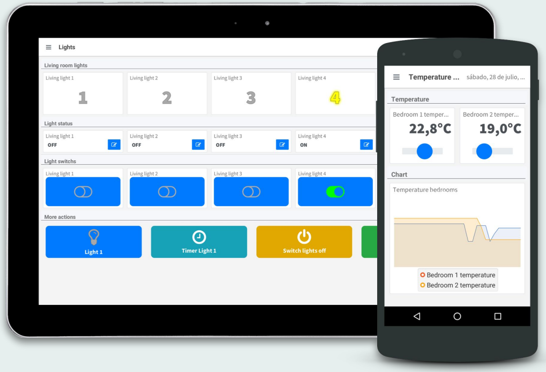
Output

Servos



Displays

Speakers



MCUs enable **TinyML**

SIZE

LOW POWER

LOW COST

MCUs enable **TinyML**

SIZE

LOW POWER

LOW COST

< 140 µW

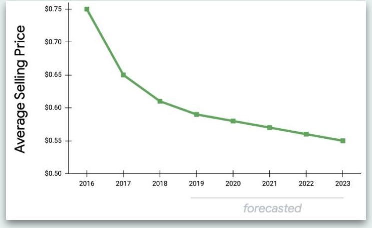
Syntian NDP100

MCUs enable **TinyML**

SIZE

LOW POWER

LOW COST



What Makes **TinyML**?

1. MCU is the building block of **TinyML** devices.
2. Theses devices are going to be pervasive or ubiquitous
3. Are they capable of running ML models?