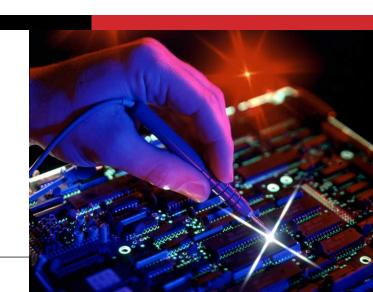


Robotics

INTRODUCTION TO TINYML

Dennis A. N. Gookyi





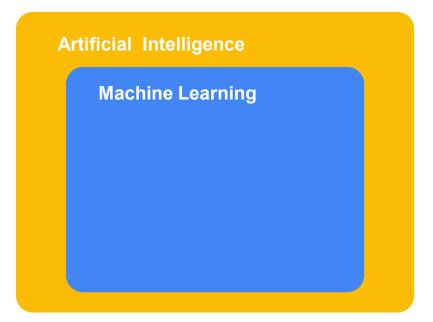
Introduction TO TinyML





MACHINE LEARNING

- Machine Learning
 - Machine Learning is a subfield of Artificial Intelligence focused on developing algorithms that learn to solve problems by analyzing data for patterns

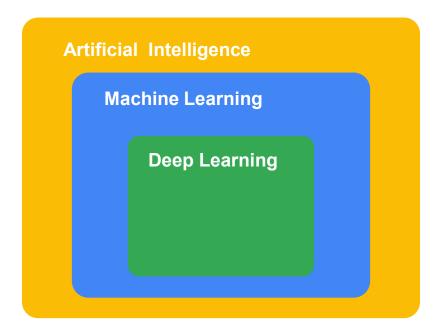






MACHINE LEARNING

- Deep Learning
 - Deep Learning is a type of Machine Learning that leverages
 Neural Networks and Big Data







Applications of machine learning



















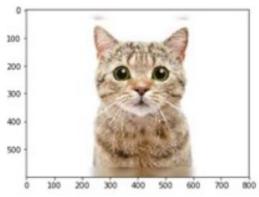




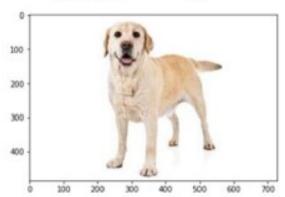
Image classification

[PREDICTION] [Prob]

Egyptian cat : 64%
tabby : 14%
bucket : 3%



[PREDICTION] [Prob]
Labrador retriever : 83%
golden retriever : 13%
bloodhound : 6%



[PREDICTION] [Prob]

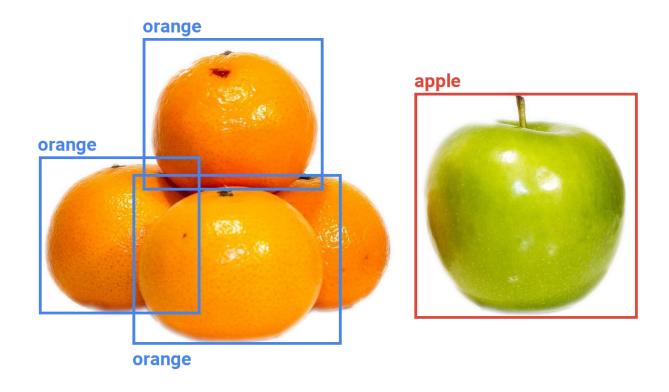
German shepherd : 60% dhole : 16% malinois : 7%







Object detection







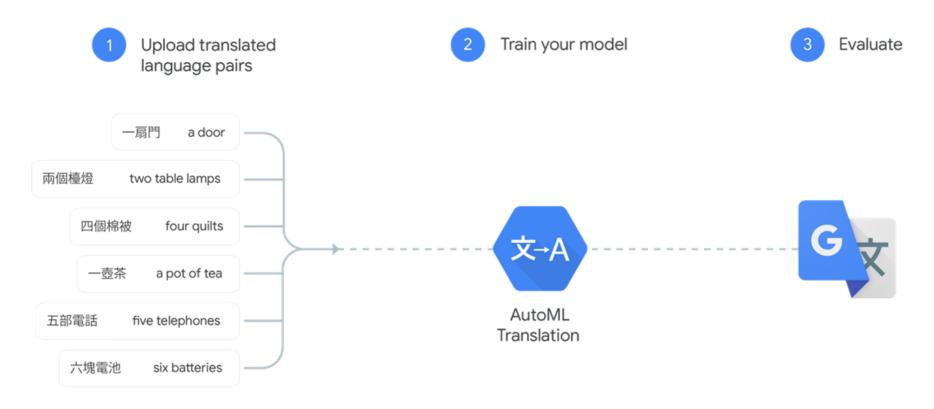
Segmentation







Machine translation

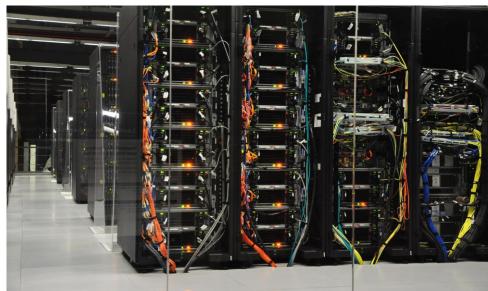






Data centers

All the capabilities on previous examples, required a remarkable amount of horsepower and computing capabilities, so what companies are doing, they are taking all these computers and jam packing them into data centers, that are all just being dedicated in order to provide machine learning capabilities today









TPUs/GPUs

- In order to be able to provide ML capability, companies like
 Google are building TPUs (Tensor Processing Units) and NVIDIA
 GPUs (Graphics Processing Units)
- Both of these computing systems are capable of running machine learning extremely fast

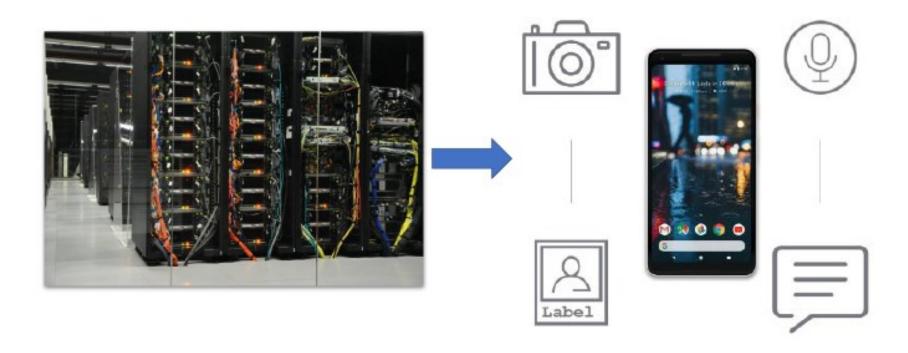








- Bigger is not always better
 - Because we can not have a Datacenter to do ML inside our phone







- Bigger is not always better
 - Because we can not have a Datacenter to do ML inside our phone



High power High bandwidth **High latency**





Low power
Low bandwidth
Low latency





- Bigger is not always better
 - Because we can not have a Datacenter to do ML inside our phone







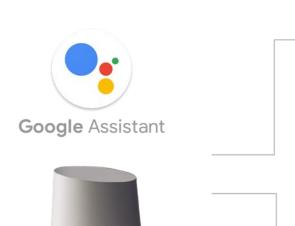
- Bigger is not always better
 - Because we can not have a Datacenter to do ML inside our phone







End devices



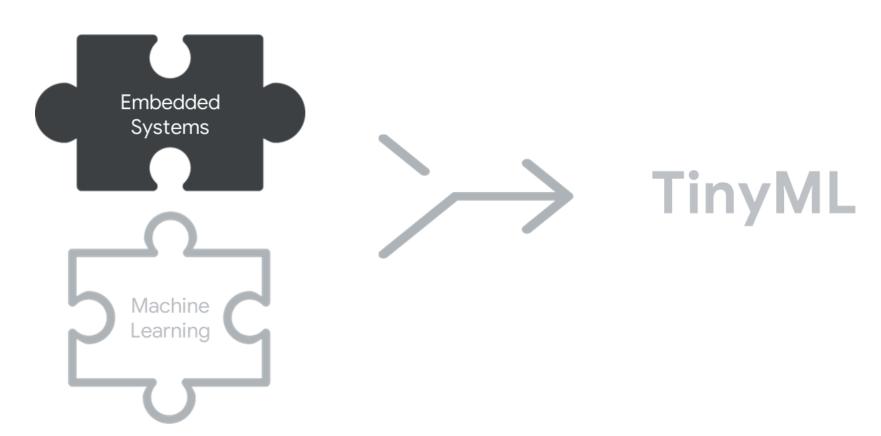








❖ What Makes TinyML?







Example





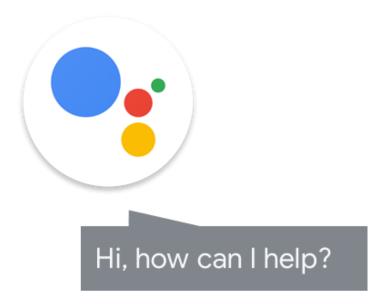
Google Assistant





Example



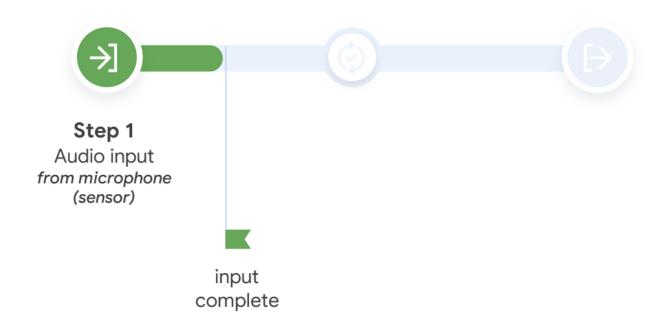






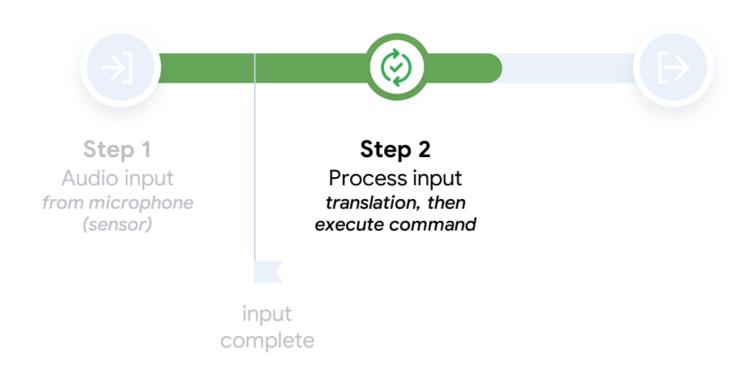






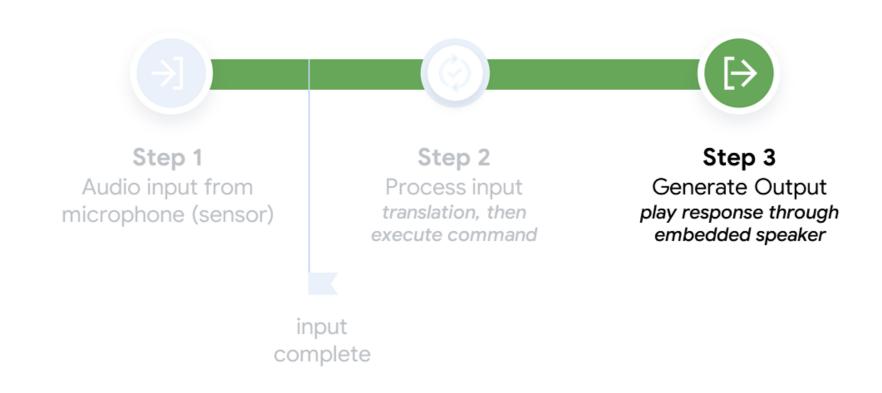








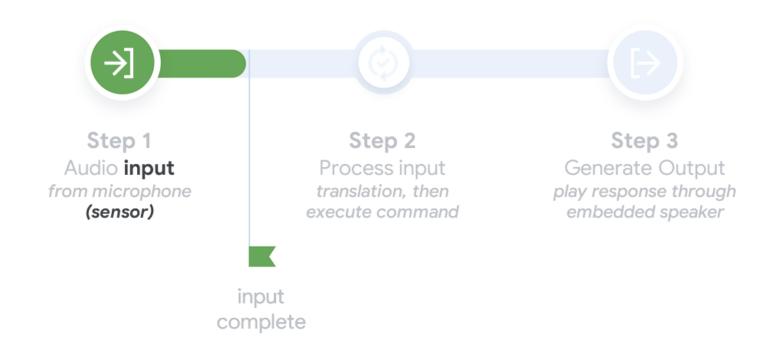








Inputs







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





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 SensorsFingerprint, Heart rate, etc.

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, 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





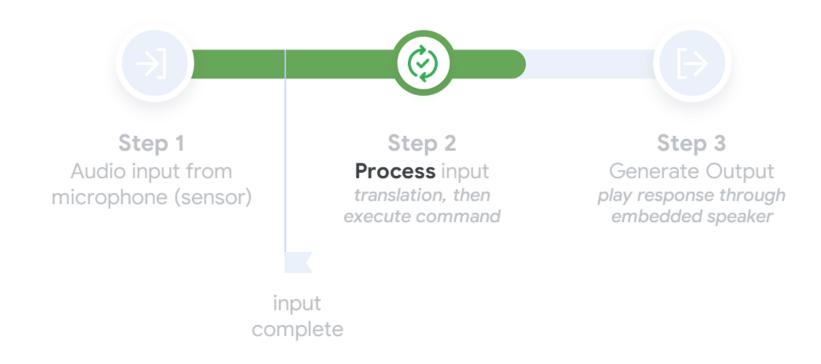
Endpoints have sensors, tons of sensors







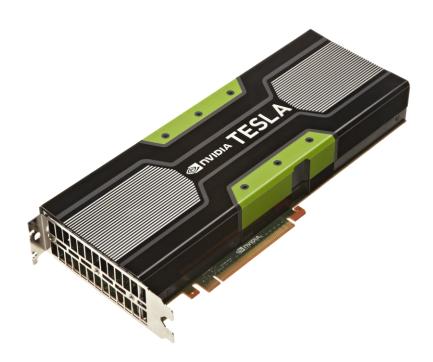
Processing







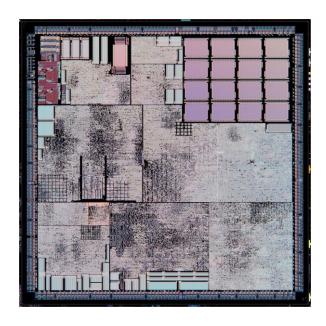
Thinking big







Thinking big







Thinking big

BIG GPU / CPU 561mm²





Thinking small

BIG GPU / CPU 561mm²







Thinking small

BIG GPU / CPU 561mm²







Thinking small

BIG GPU / CPU 561mm²

SMALL

Mobile SoC 83mm²





Thinking tiny

BIG GPU / CPU 561mm²

SMALL

Mobile SoC 83mm²







Thinking tiny

BIG GPU / CPU 561mm²

SMALL

Mobile SoC 83mm²







Thinking tiny

BIG GPU / CPU 561mm²

SMALL

Mobile SoC 83mm²







Thinking tiny

BIG GPU / CPU 561mm²

SMALL

Mobile SoC 83mm²

TINY

Apple 0778 *30mm*²





Thinking record breaking

BIG GPU / CPU 561mm²

SMALL

Mobile SoC 83mm² TINY

Apple 0778 *30mm*²

Kinetis KL03 3.2mm²





Thinking record breaking

BIG GPU / CPU 561mm² **SMALL**

Mobile SoC 83mm² TINY

Apple 0778 30mm²

world's smallest ARM-Powered MCU

48MHz, 32KB flash, 20-pin

Kinetis KL03 3.2mm²





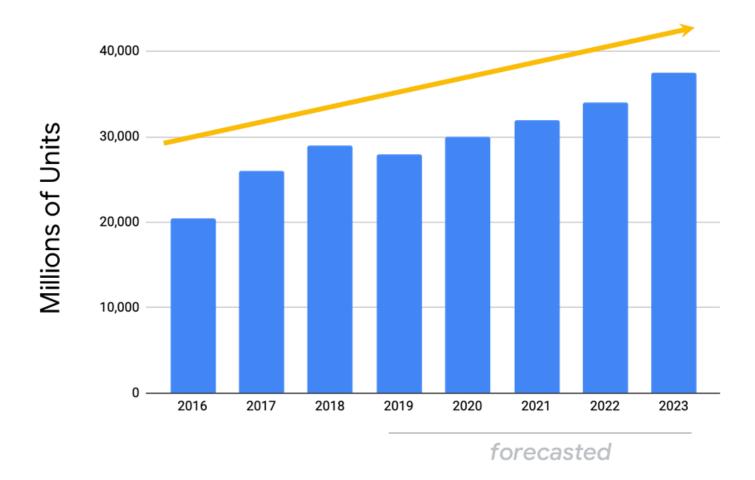
Thinking record breaking







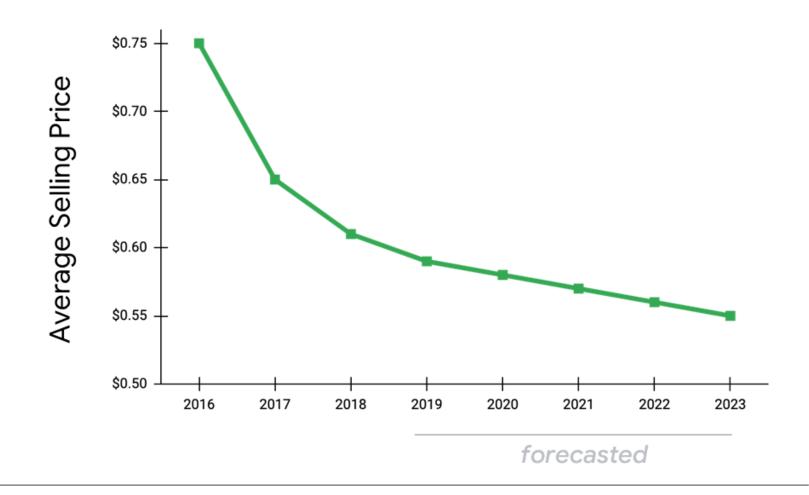
MCU demand forecast







MCU pricing forecast







Comparing power

BIG GPU / CPU

300W NVIDIA Tesla K80 **SMALL**

3.64W Apple A12

Neural Decision Processor

Always-on deep learning speech/audio recognition

Ultra low power, 128KB SRAM, 12-pin, 2.52mm²

140 μW

Syntiant NDP100





Comparing power



Use case: button cell battery

Neural Decision Processor

Always-on deep learning speech/audio recognition

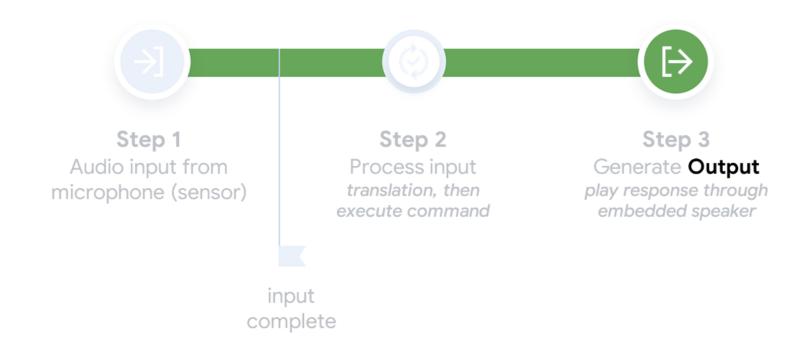
Ultra low power, 128KB SRAM, 12-pin, 2.52mm²

140 μWSyntiant NDP100





Output

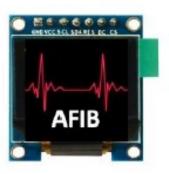


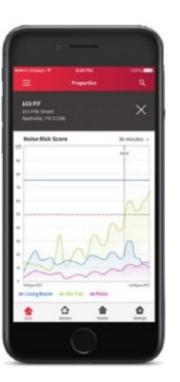




Output











❖ MCUs enable TinyML

SIZE

LOW POWER LOW COST

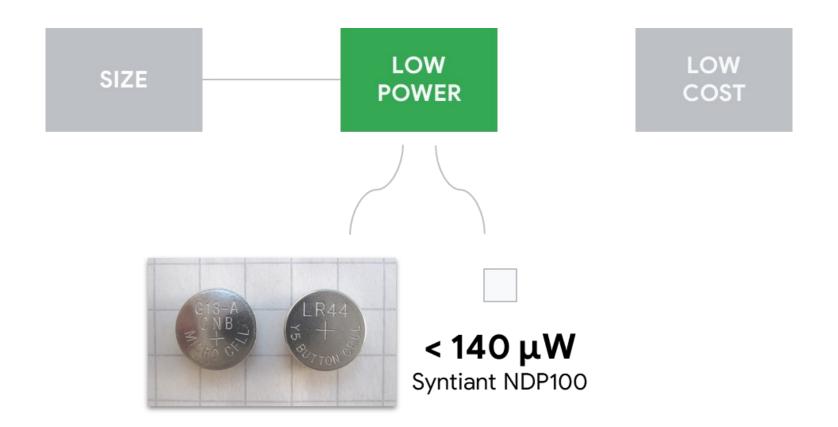








❖ MCUs enable TinyML

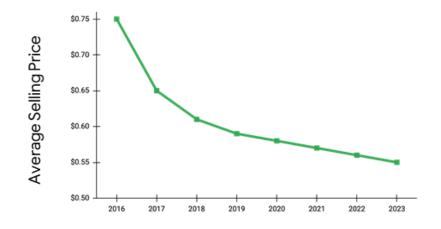






MCUs enable TinyML

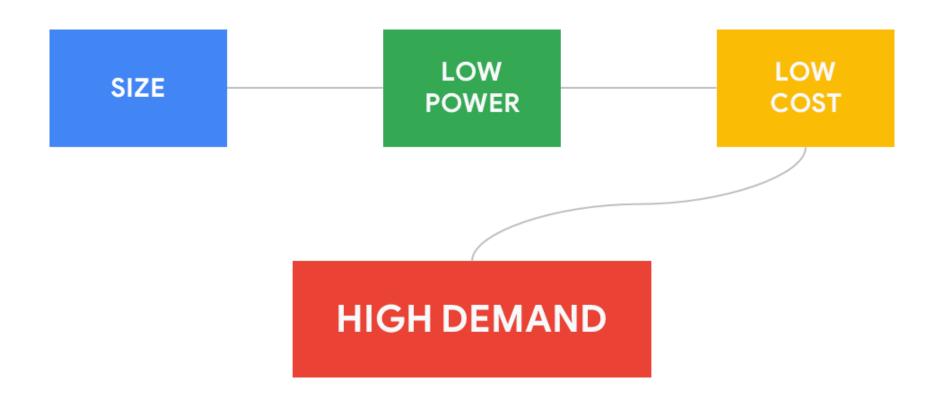








❖ MCUs enable TinyML







What makes TinyML

Embedded Systems

Machine Learning

