

## ECML PKDD 2021 Tutorial

# Machine Learning Meets Internet of Things: From Theory to Practice

## Part I: ML for IoT Devices

Bharath Sudharsan

A World Leading SFI Research Centre





**Missing Child**



**Icy Sidewalk**



**Spilled Liquid**



**Ignored Display**



**Long Line**

Valuable extractable knowledge at the edge

- Cisco annual Internet report
  - ✓ By 2023, there will be more than three times more networked devices on Earth than humans
  - ✓ About half of the global connections will be machine-to-machine (M2M) connections
- Push applications, data and computing power to the edge of the Internet
- Deploying, managing apps outside the data centers, on the customer-owned devices

# IoT Edge Computing - Hardware View

ARM Cortex-M0 MCU  
based BLE beacon



Powerful CPU + basic GPU based  
SBCs (single board computers)



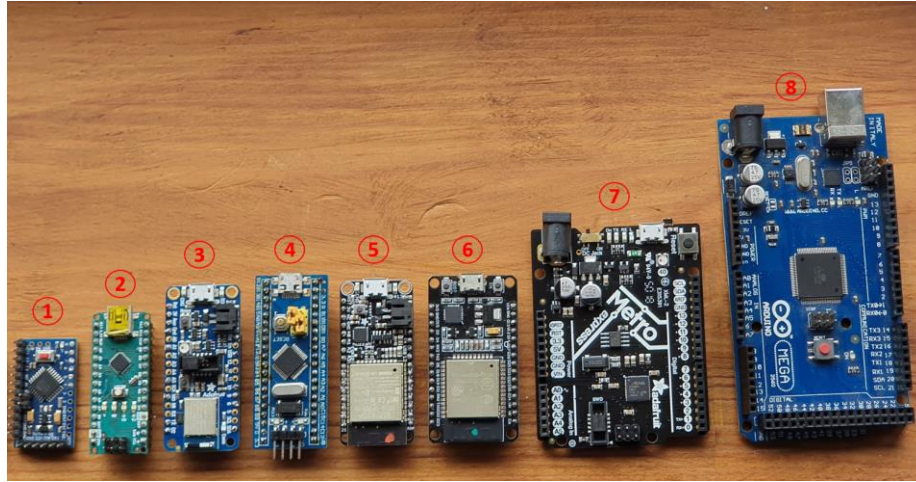
Edge gateway with  
GPUs and SSDs



Edge computing hardware: highly resource constrained -> high resource (left to right)

- MCUs and small CPUs: BLE beacons, smart bulbs, smart plugs, TV remotes, fitness bands
- SBCs: Raspberry Pis, BeagleBones, NVIDIA Jetsons, Latte Pandas, Intel NUCs, Google Coral
- GPU accelerated: AWS snowball, Digi gateways, Dell Edge Gateways for IoT, HPE Edgeline
- Why MCU? Roughly **50 billion** MCU chips were shipped in 2020 (market estimates), which far exceeds other chips like GPUs & CPUs (only 100 million units sold)

# IoT Edge Computing - Hardware Spec



Board	MCU & Board Name	Specification					
		Bits	EEP ROM	SRAM	Flash	Clock (MHz)	FP
#1	ATmega328P Arduino Pro Mini, Nano	8	1kB	2kB	32kB	16	✗
#3	nRF52840 Adafruit Feather	32	-	256kB	1MB	64	✓
#4	STM32f103c8 Blue Pill	32	-	20kB	128kB	72	✗
#5	Adafruit HUZZAH32, Generic ESP32	32	-	520kB	4MB	240	✓
#7	ATSAMD21G18 Adafruit METRO	32	-	32kB	256kB	48	✗
#8	ATmega2560 Arduino Mega	8	4kB	8kB	256kB	16	✗

Popular open-source MCU boards with specification

- No file system, Lack Python support (only C, C++), only few hundred MHz clock speed, low SRAM and Flash memory
- No parallel execution units, lack inbuilt hardware accelerators such as APU (Accelerated Processing Unit), KPU (convolution operation accelerator), FPU (Floating-point accelerator), and FFT (Fourier transform accelerator)

# TensorFlow Lite for Microcontrollers

Confirm  
Smart Manufacturing

Train a  
model

Convert the  
model

Optimize  
the model

Deploy the  
model at  
Edge

Make  
inferences  
at Edge

 TensorFlow Lite



**Command:**

```
xxd -i  
converted_  
quantised_  
model file >  
translated c  
byte array  
of model.cc
```

**Translated C byte array:**

```
unsigned char  
converted_quantised_model[] = {  
  0x18, 0x00, 0x00, 0x00,  
  0x54, 0x46, 0x4c, 0x33,  
  0x00, 0x00, 0x0e, 0x00,  
  ...  
  ...  
};  
unsigned int  
converted_quantised_model_len = 21200;
```

Android

Linux Embedded  
devices – Raspberry Pi

iOS

Microcontrollers

Model initialization  
from c byte array

Initialize Reporter &  
Perform Unit Test

Declare Ops  
Provider to access  
model operations

ip, op & intm arrays  
memory pre-  
allocation

Interpreter  
initialization for  
variables passing

Model's tensors  
memory allocation

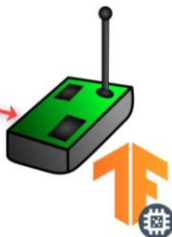
ip shape validation

Real-time data  
provision to model

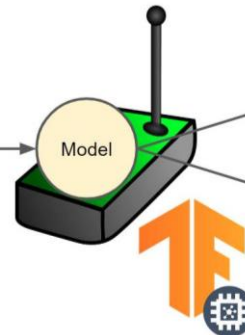
Model Execution

Obtain model's  
output tensor

Loop Inference



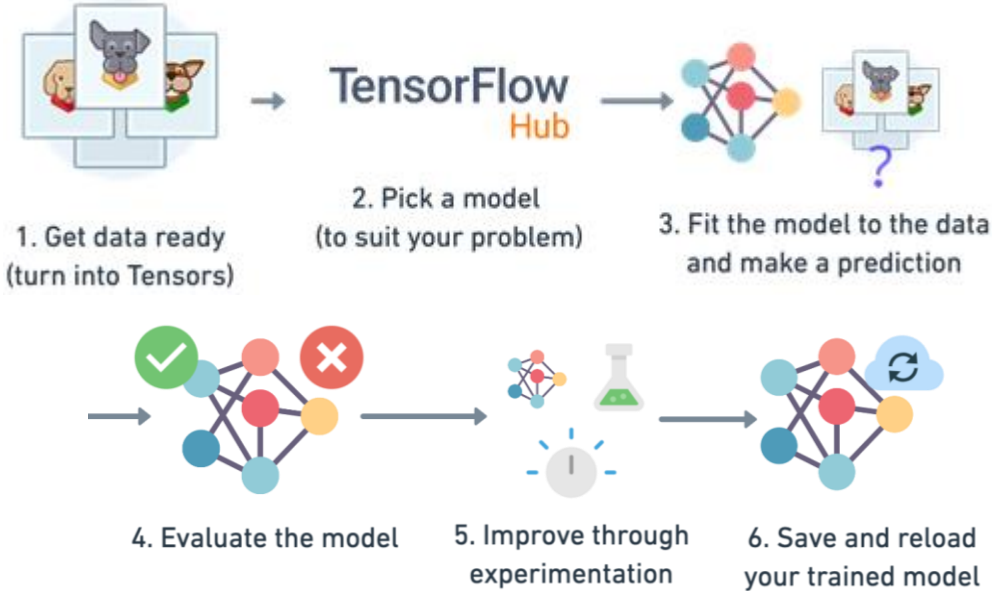
Unseen  
data



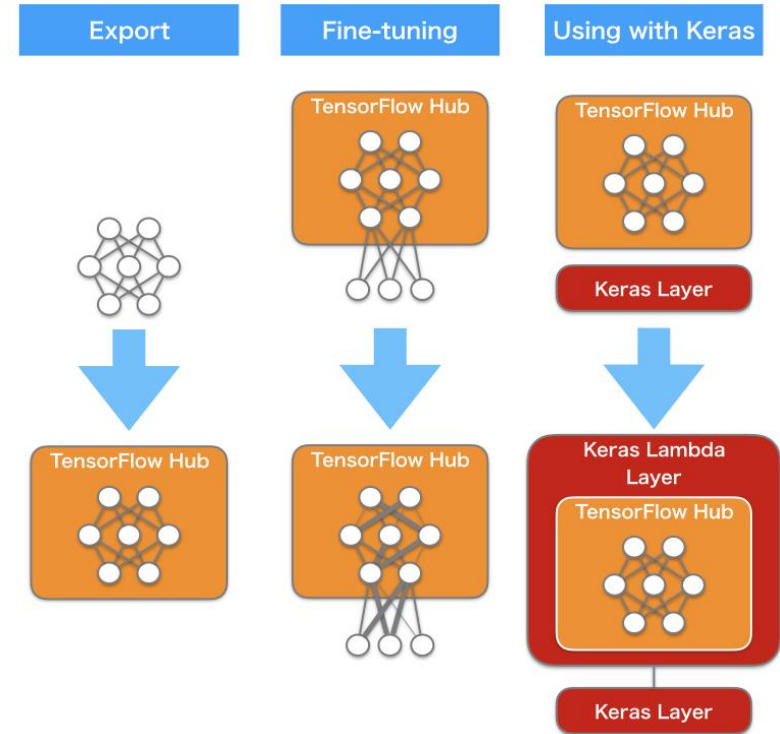
Cat

Not cat

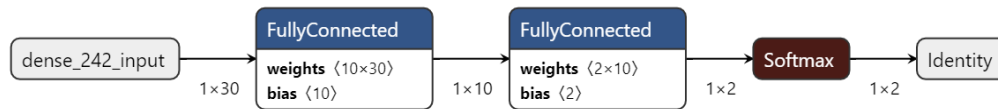




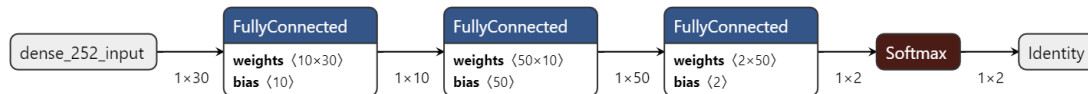
- Models ready to use from off the shelf
- Optimizers applicable: Quantization, Pruning, etc.
- Can add additional tunable layers at the network end



- Demo: Download popular models and visualize and brief analysis using <https://netron.app/>
  - ✓ yolo-v5-tflite from [https://tfhub.dev/neso613/lite-model/yolo-v5-tflite/tflite\\_model/1](https://tfhub.dev/neso613/lite-model/yolo-v5-tflite/tflite_model/1)
  - ✓ 3 types of fully connected Neural Networks (NNs) from <https://github.com/bharathsudharsan/TinyML-Benchmark-NNs-on-MCUs>



FC 1 x 10: 1 layer with 10 neurons

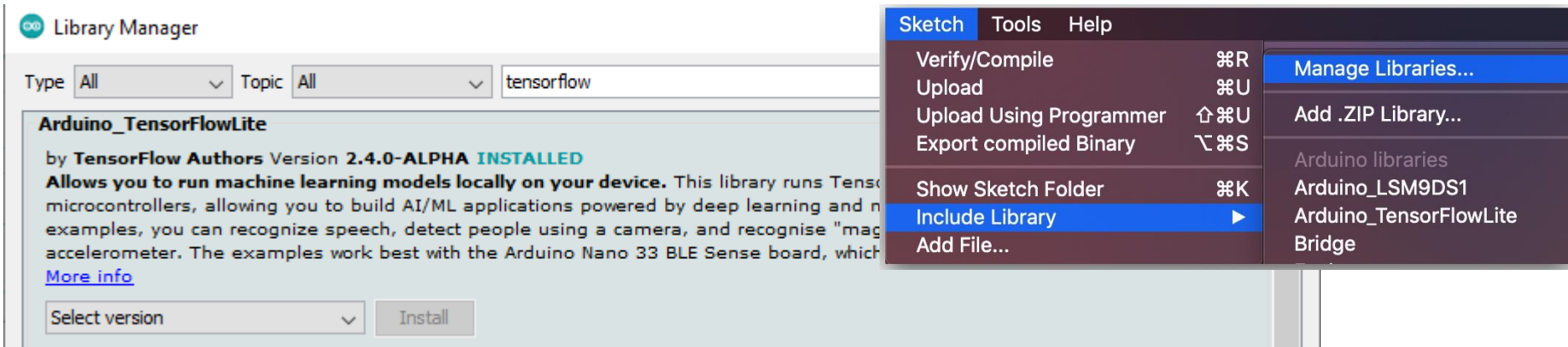


FC 10 + 50: 2 layers, where 1st layer contains 10 neurons, and 2nd layer contains 50 neurons



FC 10 x 10: 10 layers, where each layer contains 10 neurons

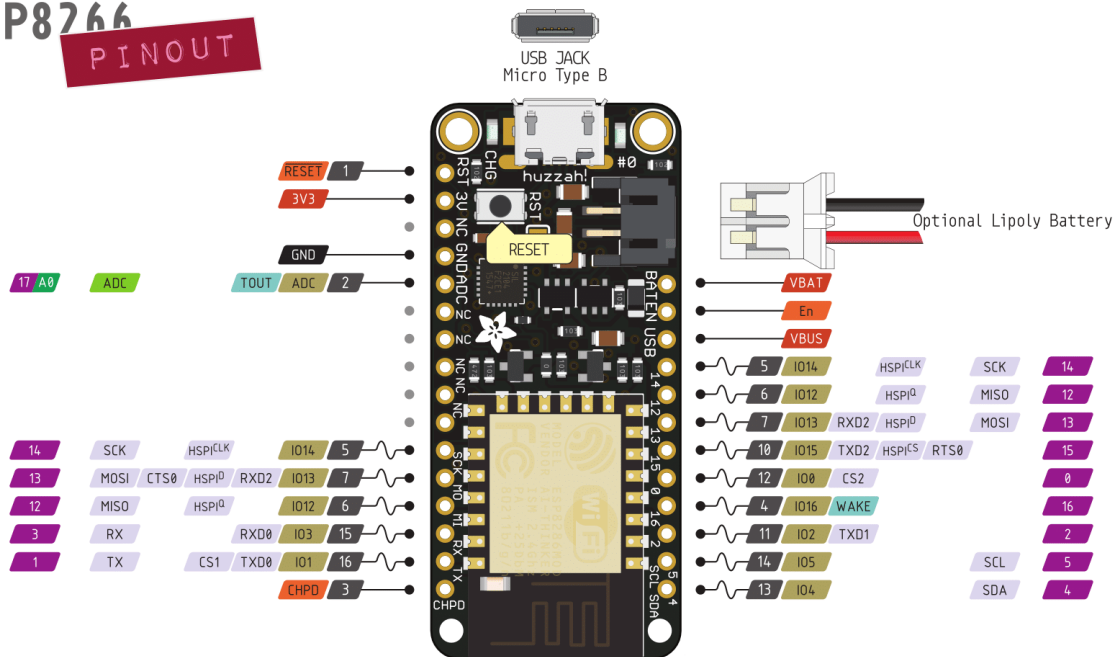
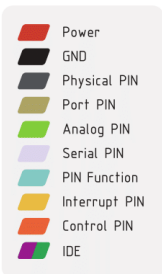
# Arduino IDE + TF Lite Library



- Install the Arduino TensorFlow Lite library
- Run the Arduino TF Lite hello world example
  - ✓ Files -> Examples -> Arduino\_TensorFlowLite -> hello\_world. The example sketch should appear
  - ✓ Upload on MCU, open the Serial Plotter, observe LED
  - ✓ Model trained to replicate sine function. Generates data pattern to either blink LEDs or control an animation



**Confirm**  
Smart Manufacturing

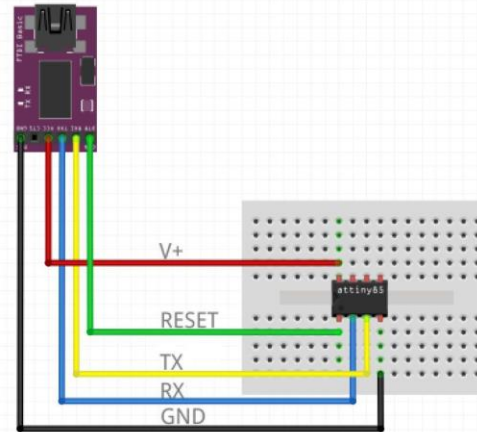
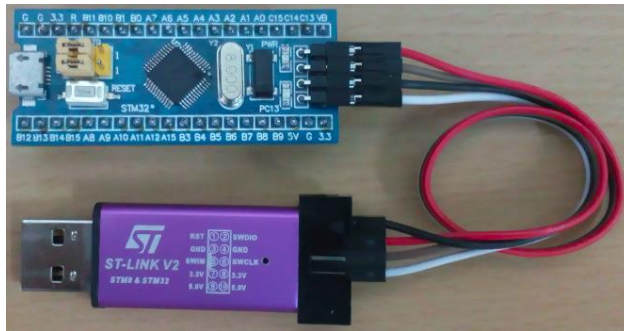
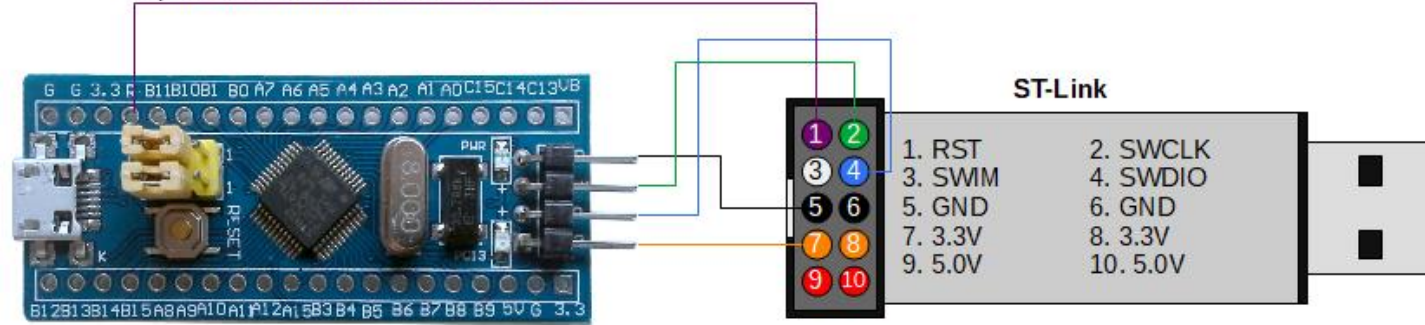


- Pinout summary
  - ✓ Power
  - ✓ Logic
  - ✓ I2C & SPI
  - ✓ GPIO
  - ✓ Analog

 PWM Pin

# Dev Boards - Hardware View

Optional: enables hardware reset mode



- ST-LINK V2 programmer provides a more reliable way to flash the Blue Pill boards
- Full range of STM32 SWD interface debugging
- The tiniest Arduino ATtiny85 is flashed (programmed burned) using UART IC

# Dev Boards - ML Applications



## Stator Yoke Temperature

69

Predict the stator temperature of an electric car with a thermal sensor



## Is my fan obstructed? STM32

99

Detect with an accelerometer if my fan is obstructed



## Lock picking

93

Detect if a lock is being open with a key or using picks

<https://cartesiam.ai/product/>



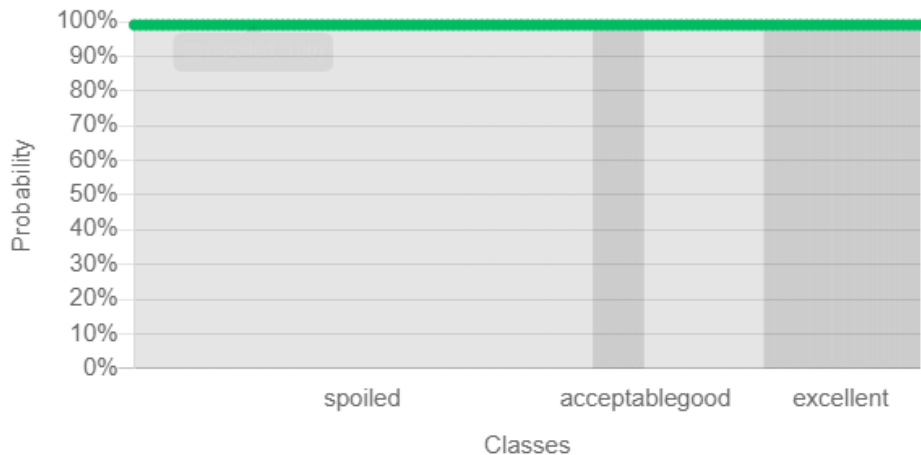
## Beef quality monitoring

100

Electronic nose for beef quality monitoring

## Result from NanoEdge™ AI Studio

NanoEdge™ AI Studio Engine used: v2021.05.04.0



ACCURACY

100.00%

CONFIDENCE

100.00%

RAM

0.1kB

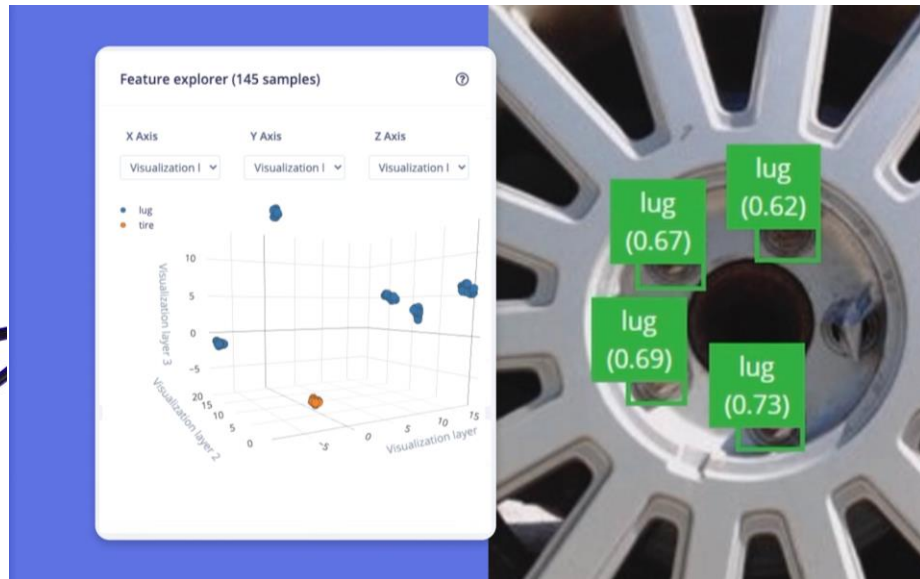
+ Buffer 0kB

FLASH

0.2kB



# Dev Boards - ML Applications



<https://www.edgeimpulse.com/blog>





- A small device with an enormous impact on our lives. For examples
  - ✓ TinyLSTMs: Efficient Neural Speech Enhancement for Hearing Aids
  - ✓ A Tiny CNN Architecture for Medical Face Mask Detection for Resource-Constrained Endpoints

# Confirm

Smart Manufacturing

Confirm  
Smart Manufacturing



Contact: Bharath Sudharsan  
Email: [bharath.sudharsan@insight-centre.org](mailto:bharath.sudharsan@insight-centre.org)

[www.confirm.ie](http://www.confirm.ie)

  
Science  
Foundation  
Ireland For what's next