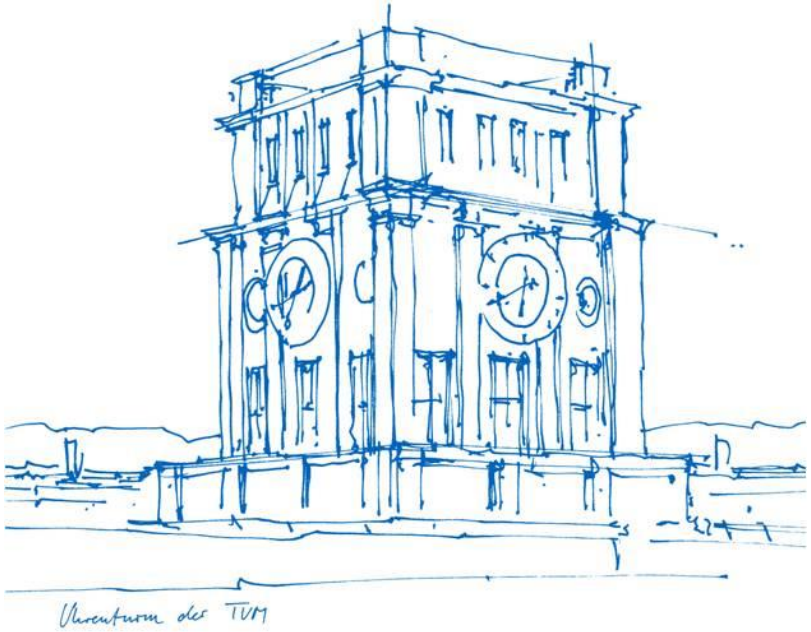
Sprint 3

Ahmed Kaddah, Shao Jie Hu Chen, Marlon Müller

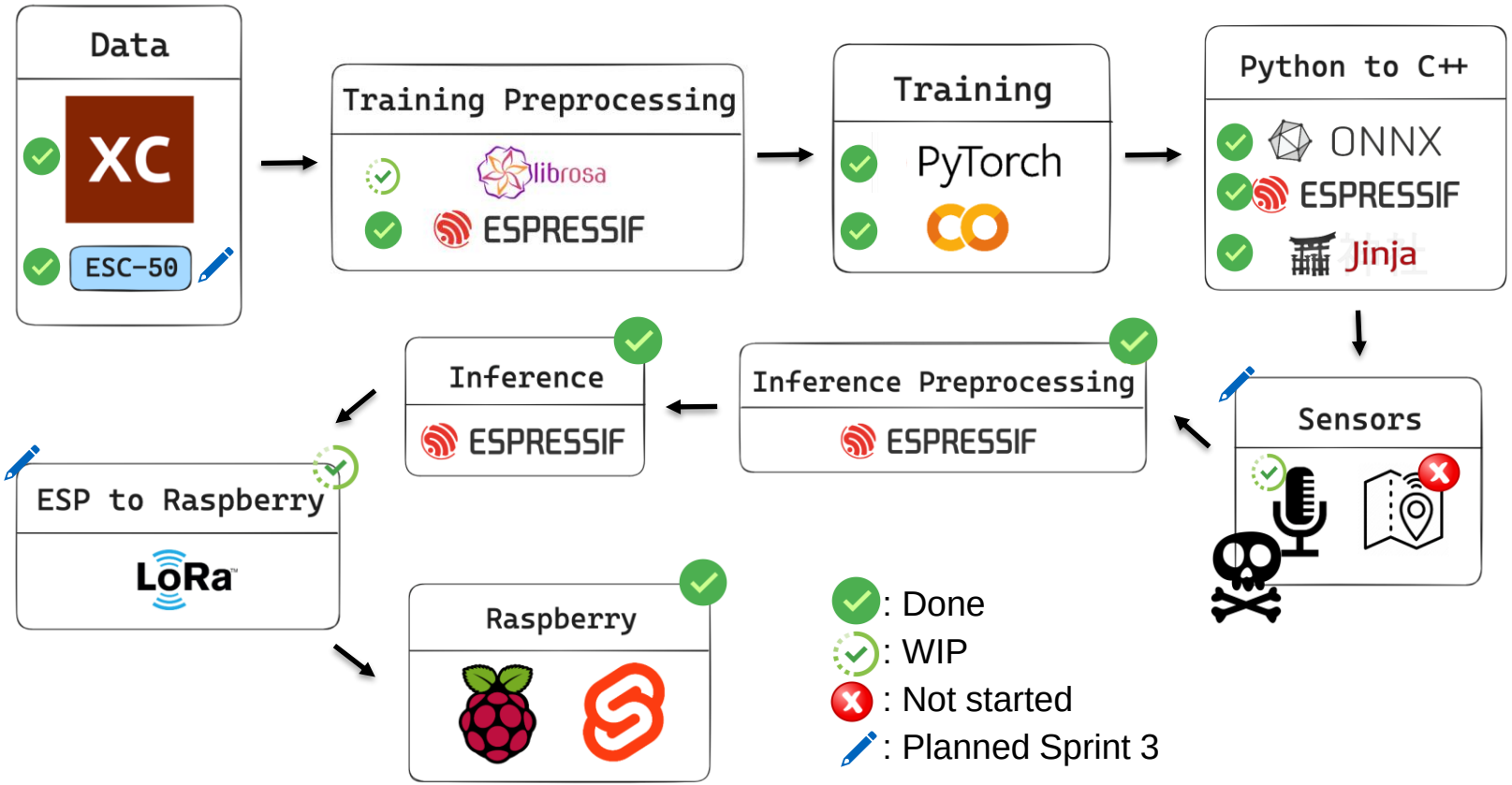
Edge Computing and the Internet of Things

Technische Universität München

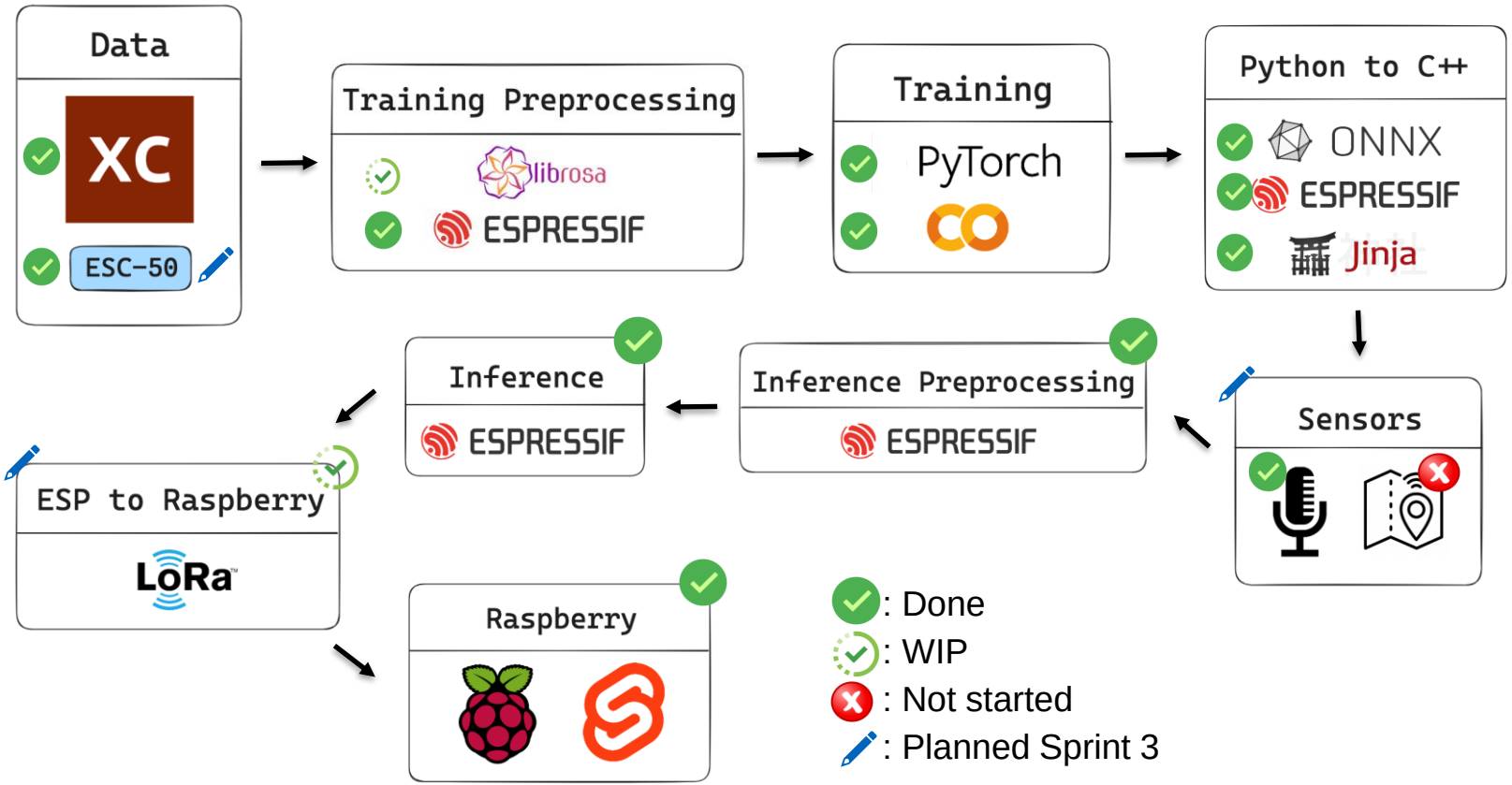
München, 12.01.2024



Overview

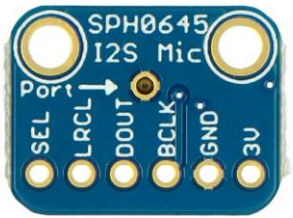

Data
XC
ESC-50
Training Preprocessing
librosa
ESPRESSIF
Training
PyTorch
Python to C++
ONNX
ESPRESSIF
Jinja
Sensors
Inference Preprocessing
ESPRESSIF
Inference
ESPRESSIF
ESP to Raspberry
LoRa
Raspberry
: Done
: WIP
: Not started
: Planned Sprint 3

Overview

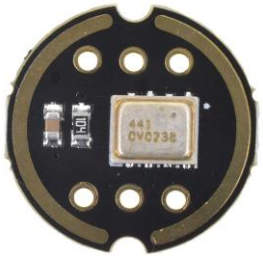

Data
XC
ESC-50
Training Preprocessing
librosa
ESPRESSIF
Training
PyTorch
Python to C++
ONNX
ESPRESSIF
Jinja
Sensors
Inference Preprocessing
ESPRESSIF
Inference
ESPRESSIF
ESP to Raspberry
LoRa
Raspberry
: Done
: WIP
: Not started
: Planned Sprint 3

Microphone

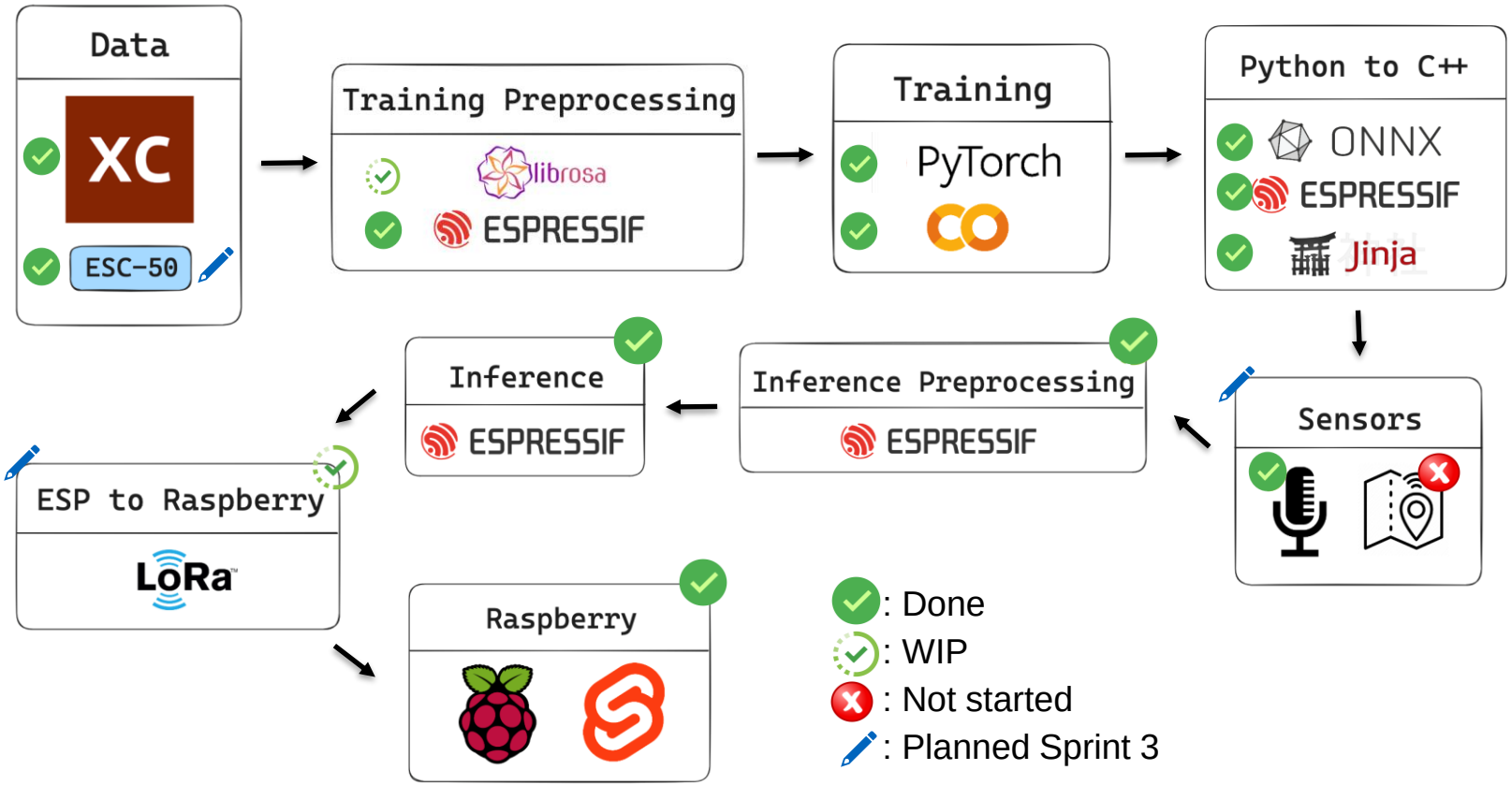
SPH0645



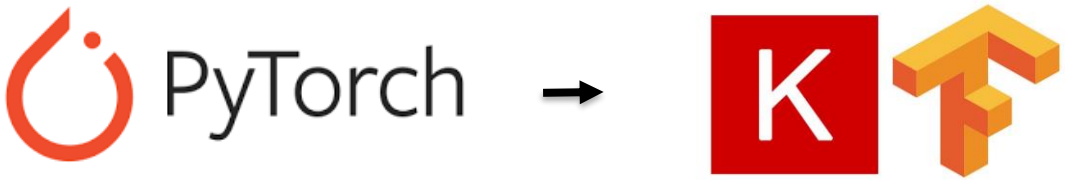
INMP411



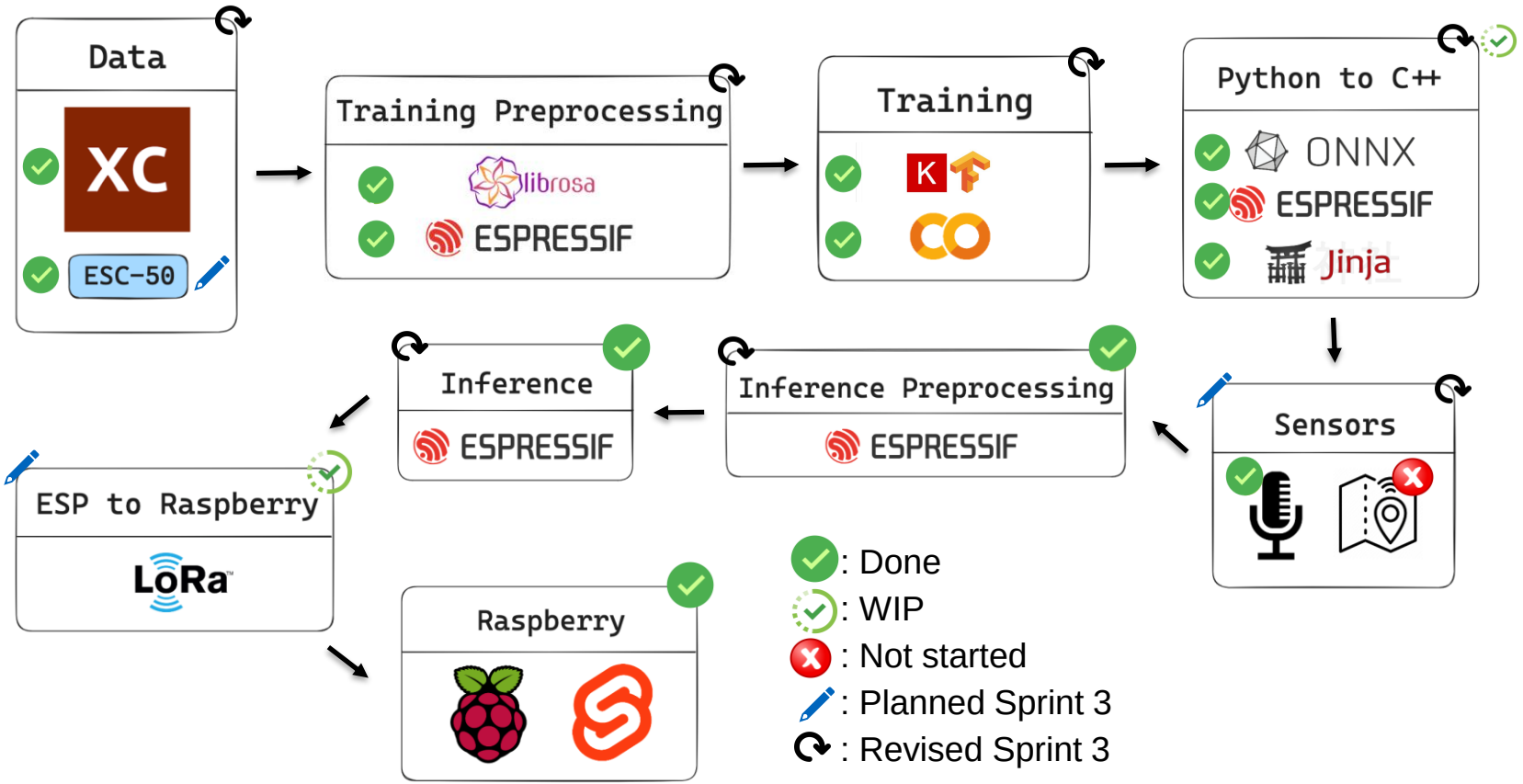
Overview


Data
XC
ESC-50
Training Preprocessing
librosa
ESPRESSIF
Training
PyTorch
Python to C++
ONNX
ESPRESSIF
Jinja
Sensors
Inference Preprocessing
ESPRESSIF
Inference
ESPRESSIF
ESP to Raspberry
LoRa
Raspberry
: Done
: WIP
: Not started
: Planned Sprint 3

PyTorch: Sunk Cost

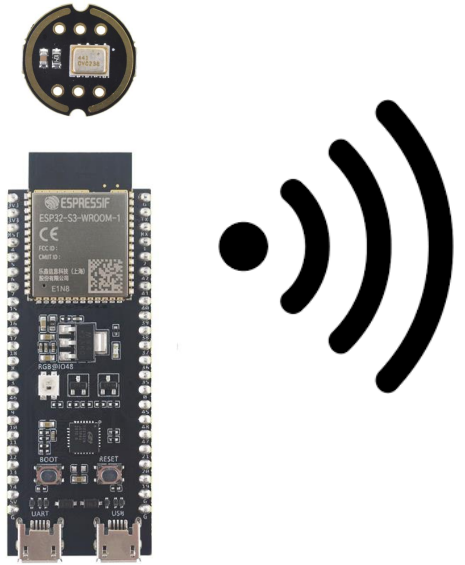


Overview


Data
XC
ESC-50
Training Preprocessing
librosa
ESPRESSIF
Training
Python to C++
ONNX
ESPRESSIF
Jinja
Sensors
Inference Preprocessing
ESPRESSIF
Inference
ESPRESSIF
ESP to Raspberry
LoRa
Raspberry
: Done
: WIP
: Not started
: Planned Sprint 3
: Revised Sprint 3

Debug: WiFi & Websocket (+ TCP/UDP)

```c
void websocket_send(esp_websocket_client_handle_t client, char *data, size_t len)
{
    if (esp_websocket_client_is_connected(client))
    {
        while (len > 0)
        {
            size_t chunk = len > CHUNCK_SIZE ? CHUNCK_SIZE : len;
            esp_websocket_client_send_bin(client, data, chunk, portMAX_DELAY);
            data += chunk;
            len -= chunk;
        }
    }
}
```



Training

- Audio samples: ≈ 2300
- Classes: 4
- Audio length: 5s
- Sampling rate: 16kHz
- 32 MFCCs

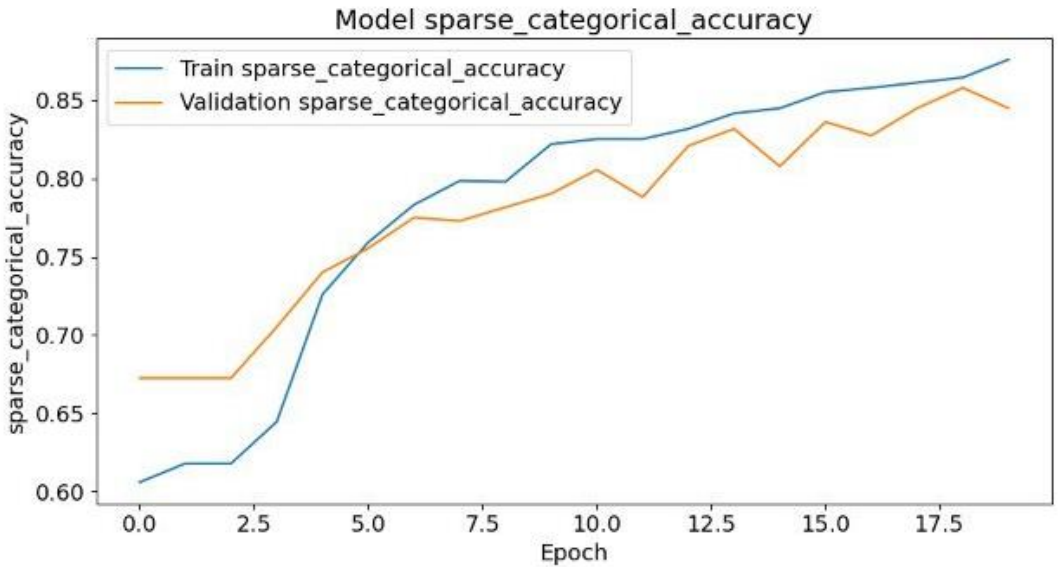
Training

- Audio samples: ≈ 2300
- Classes: 4
- Audio length: 5s
- Sampling rate: 16kHz
- 32 MFCCs

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 102, 30, 16)	256
max_pooling2d (MaxPooling2D)	(None, 34, 30, 16)	0
conv2d_1 (Conv2D)	(None, 32, 28, 16)	2320
max_pooling2d_1 (MaxPooling2D)	(None, 16, 14, 16)	0
conv2d_2 (Conv2D)	(None, 14, 12, 8)	1160
max_pooling2d_2 (MaxPooling2D)	(None, 7, 6, 8)	0
flatten (Flatten)	(None, 336)	0
dense (Dense)	(None, 64)	21568
dense_1 (Dense)	(None, 4)	260

Total params: 25564 (99.86 KB)
Trainable params: 25564 (99.86 KB)
Non-trainable params: 0 (0.00 Byte)

Training



Training

Predicted

True

```
[[ 36  35  20 187]
 [ 36  22  16 148]
 [ 24  24  13 139]
 [136 135  83 778]]
```

Validation

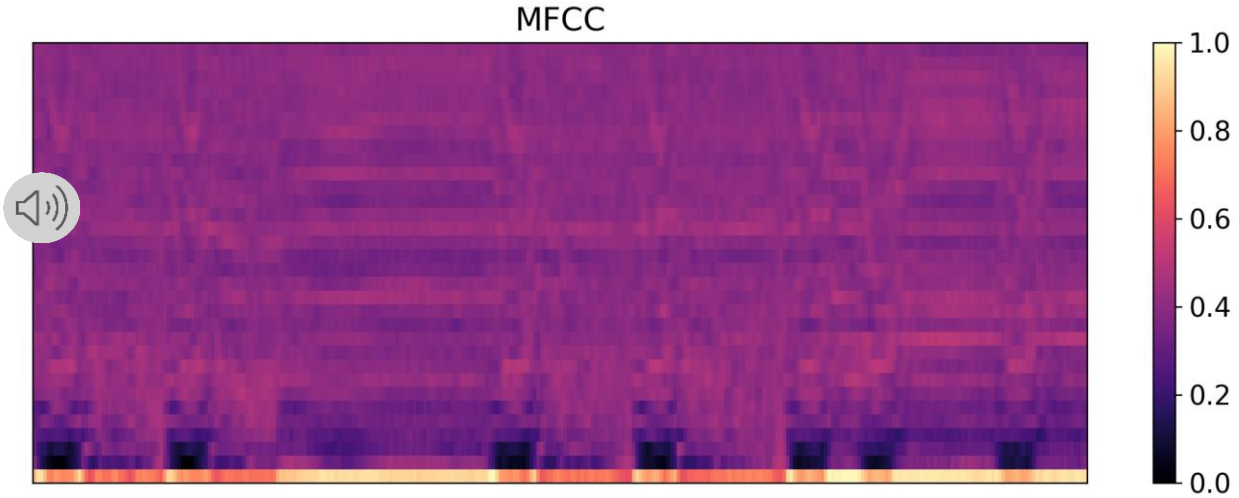
Prediced

True

```
[[ 41   0   2  19]
 [  1  31   3   7]
 [  7   2  20  17]
 [  7   0   4 297]]
```

Performance Bottlenecks

Example: Training data



Performance Bottlenecks

Example: Training data

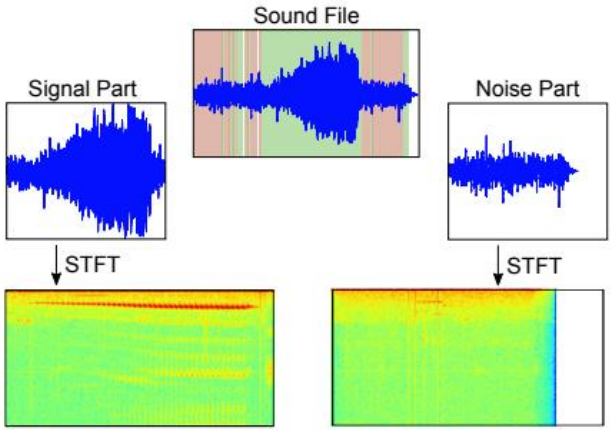
Performance issues also arise due to factors such as soundscapes, voltage spikes ...

Solution?

Generally, solutions are known, but require time to implement ...

Example: Training data

Salient audio segments



[Sprengel, Elias, et al. "Audio based bird species identification using deep learning techniques]

LoRa package format

FIELDS

- [0, 47]. 48 bits reserved for ESP ID.
- [48, 59]. 12 bits reserved for year as base 2 number.
- [60, 63]. 4 bits reserved for month as base 2 number.
- [64, 68]. 5 bits reserved for day as base 2 number.
- [69, 73]. 5 bits reserved for hour as base 2 number.
- [74, 79]. 6 bits reserved for minute as base 2 number.
- [80, 85]. 6 bits reserved for second as base 2 number.
- [86, 87]. 2 bits reserved for neural network output.
- [88, 95]. 8 bits reserved for checksum.

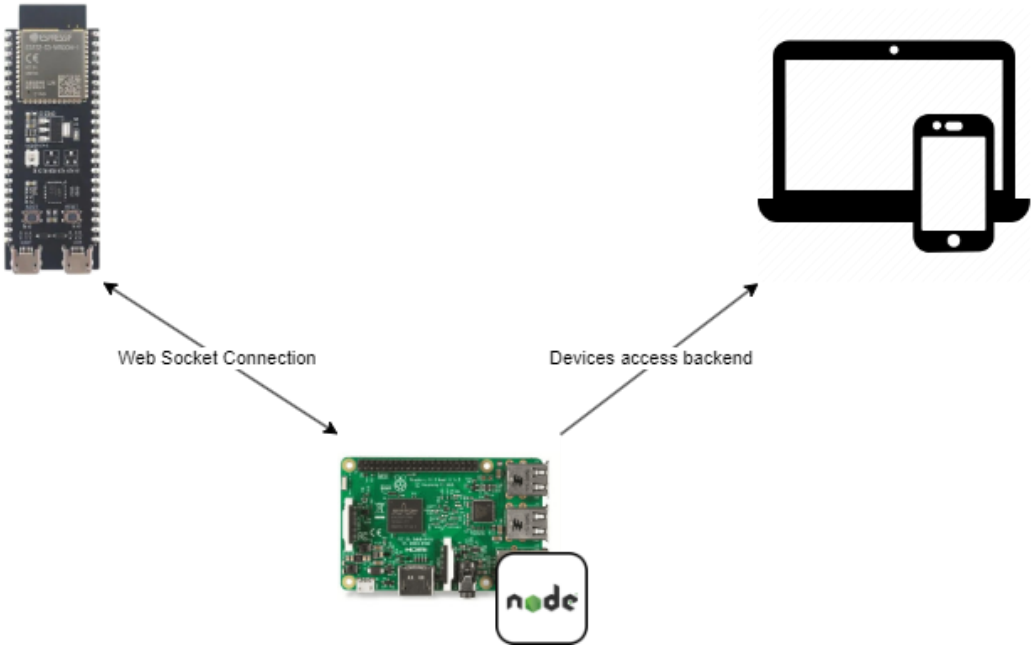
TRANSMITTED DATA

- [0, 7]. ESP-ID.
- [8, 15]. ESP-ID.
- [16, 23]. ESP-ID.
- [24, 31]. ESP-ID.
- [32, 39]. ESP-ID.
- [40, 47]. ESP-ID.
- [48, 55]. 8 first bits for Year.
- [56, 63]. 4 list bits for Year + 4 first bits Month.
- [64, 71]. 5 list bits for day + 3 first bits hour.
- [72, 79]. 2 last bits for hours + 6 bits minutes.
- [80, 87]. 6 bits seconds + 2 bits output.
- [88, 95]. 8 bits checksum.

LoRa duty cycle (12 bytes/package)

	DR6	DR5	DR4	DR3	DR2	DR1	DR0
data rate	SF7 BW 250	SF7 BW 125	SF8 BW 125	SF9 BW 125	SF10 BW 125	SF11 BW 125	SF12 BW 125
airtime	20.6 ms	41.2 ms	82.4 ms	144.4 ms	288.8 ms	577.5 ms	1,155.1 ms
1% max duty cycle	2.1 sec 1,746 msg/hour	4.1 sec 873 msg/hour	8.2 sec 436 msg/hour	14.4 sec 249 msg/hour	28.9 sec 124 msg/hour	57.8 sec 62 msg/hour	115.5 sec 31 msg/hour
fair access policy	59.4 sec (avg) **60.7 avg/hour** 1,455 msg/24h	118.7 sec (avg) **30.3 avg/hour** 727 msg/24h	237.4 sec (avg) **15.2 avg/hour** 363 msg/24h	415.8 sec (avg) **8.7 avg/hour** 207 msg/24h	831.7 sec (avg) **4.3 avg/hour** 103 msg/24h	1,663.3 sec (avg) **2.2 avg/hour** 51 msg/24h	3,326.6 sec (avg) **1.1 avg/hour** 25 msg/24h

Web socket audio stream



Sprint 4 Plan

Essential

- Integrate LoRa in RPi part with server
- Do measurements, e.g., memory or power consumption

Nice to have

- Improve DL Network
- Integrate GPS in ESP
- Light sleep for ESP
- Display location on map of classified birds on Dashboard

END