

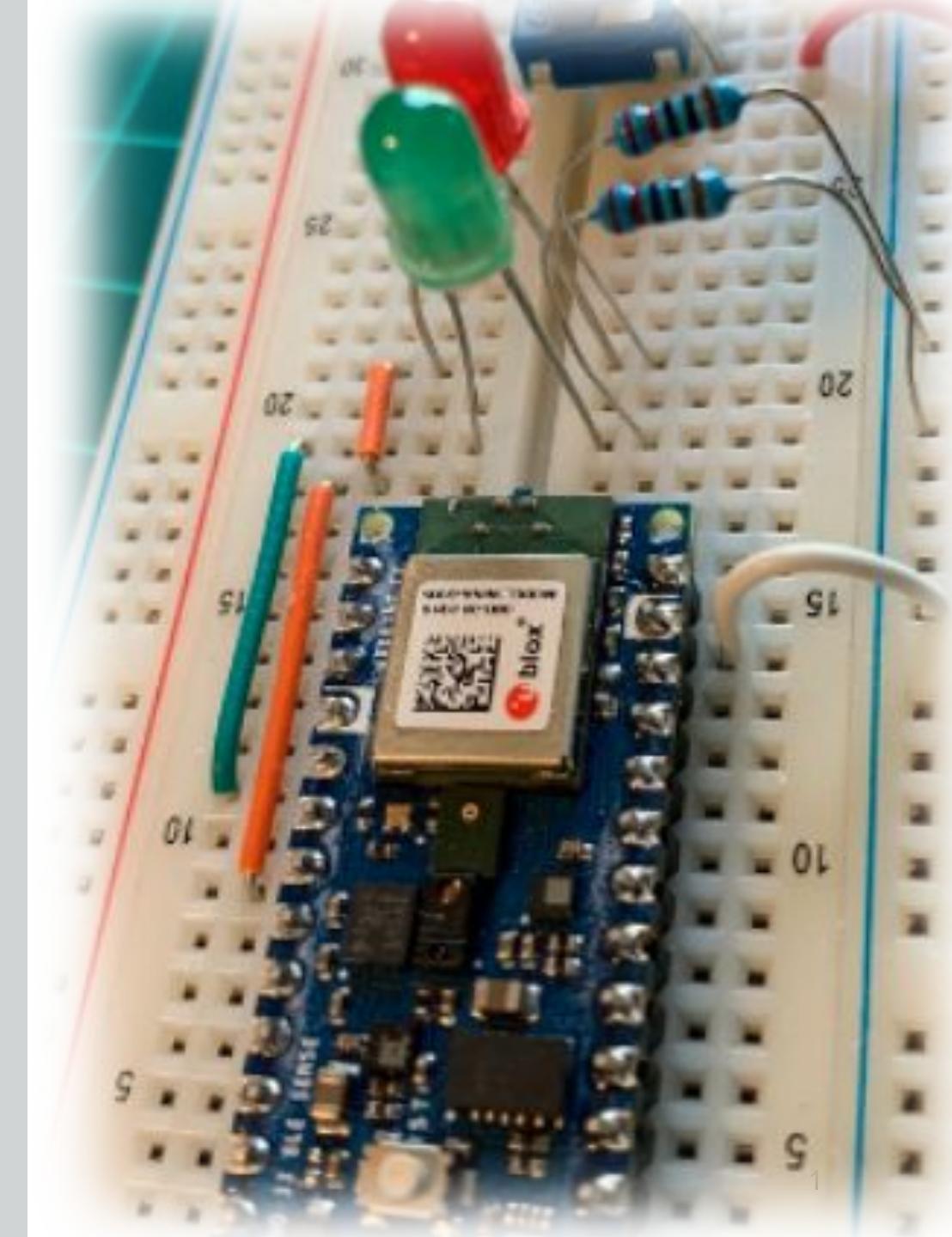
IESTI01 – TinyML

Embedded Machine Learning

12. Introduction to Edge Impulse – CNN with Cifar-10



Prof. Marcelo Rovai
UNIFEI

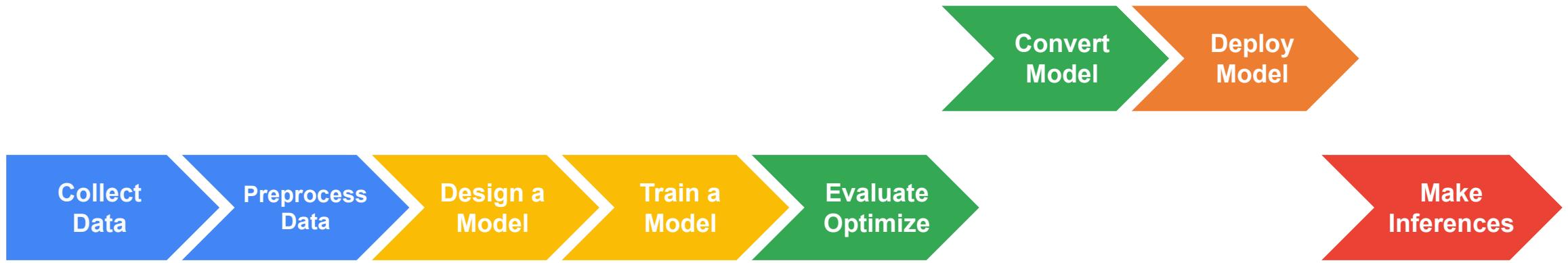


Embedded Machine Learning (TinyML) Workflow Review

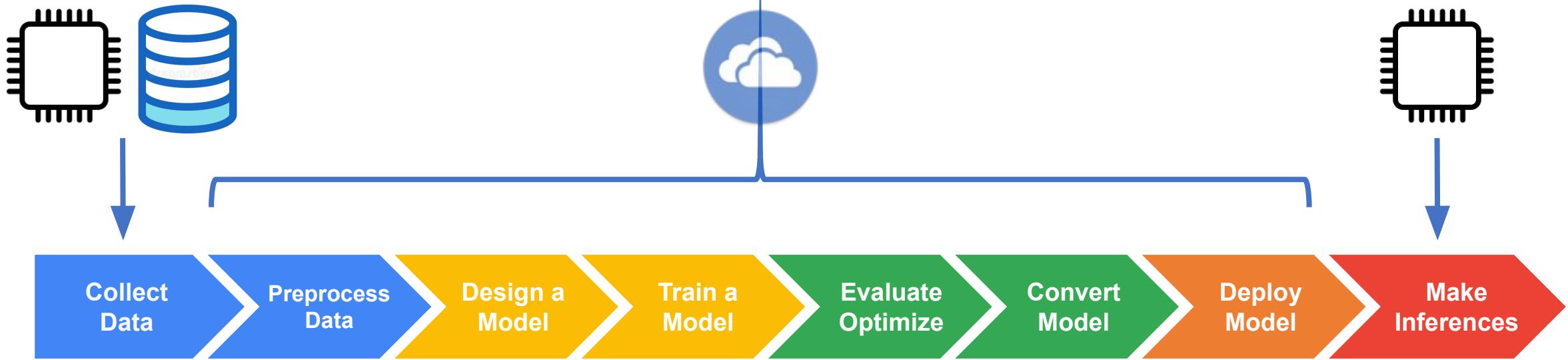
Machine Learning Workflow



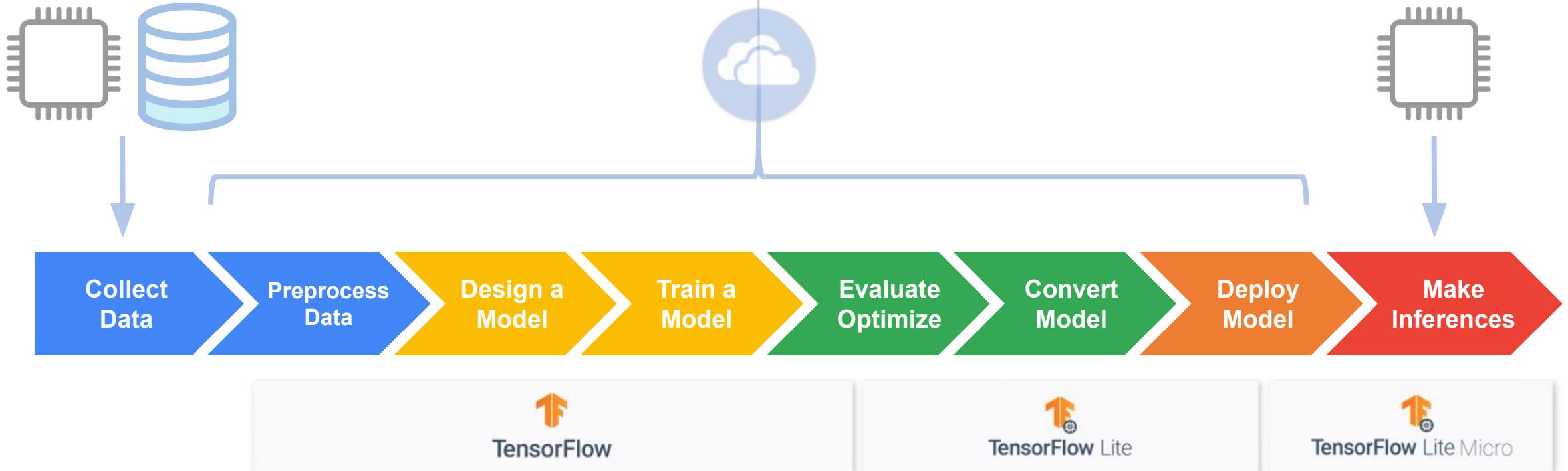
Tiny Machine Learning Workflow (“What”)



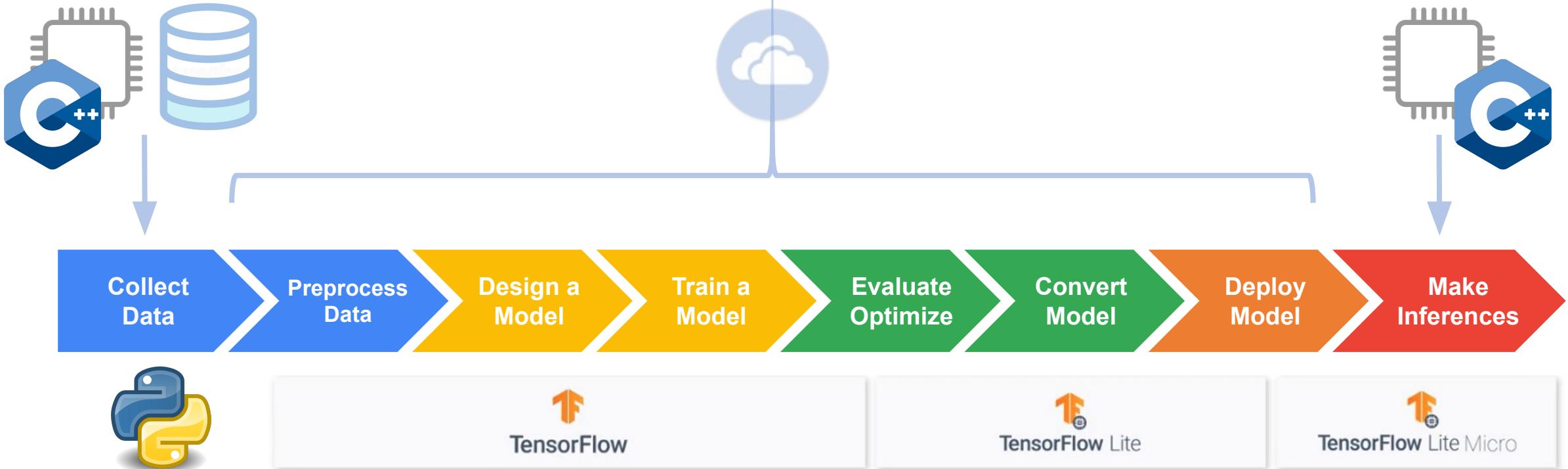
Tiny Machine Learning Workflow (“Where”)



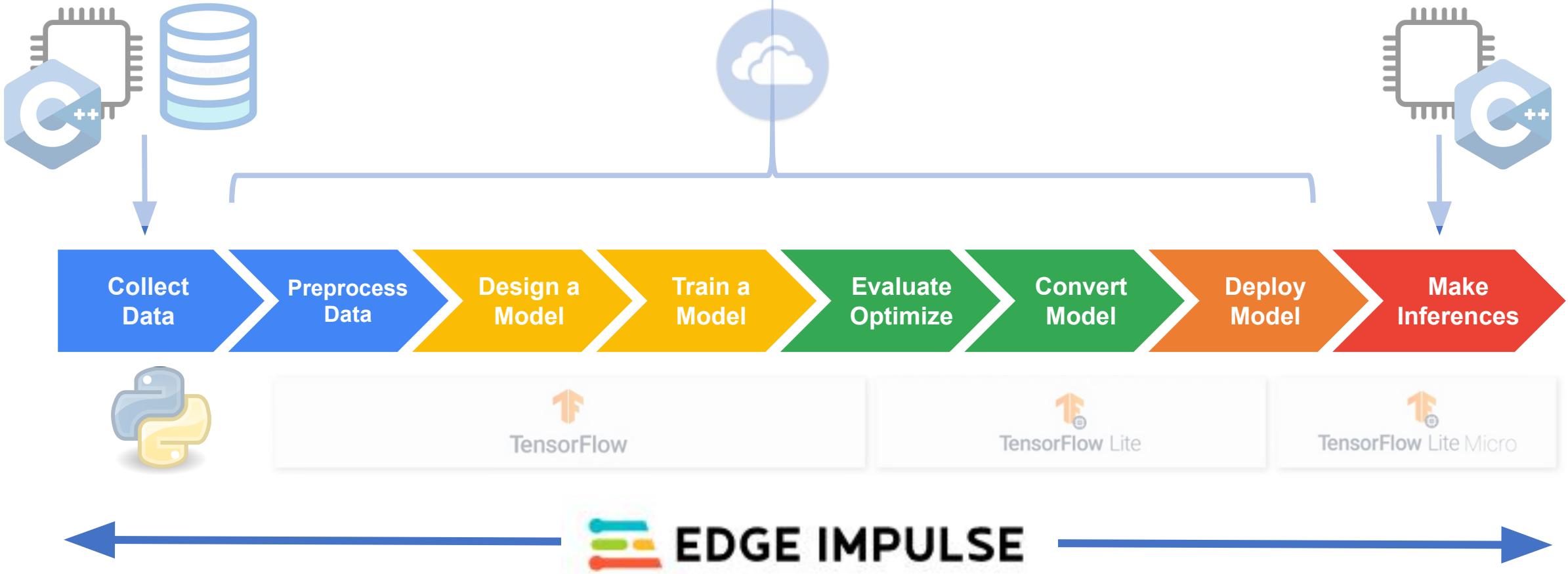
Machine Learning Workflow (“How”)

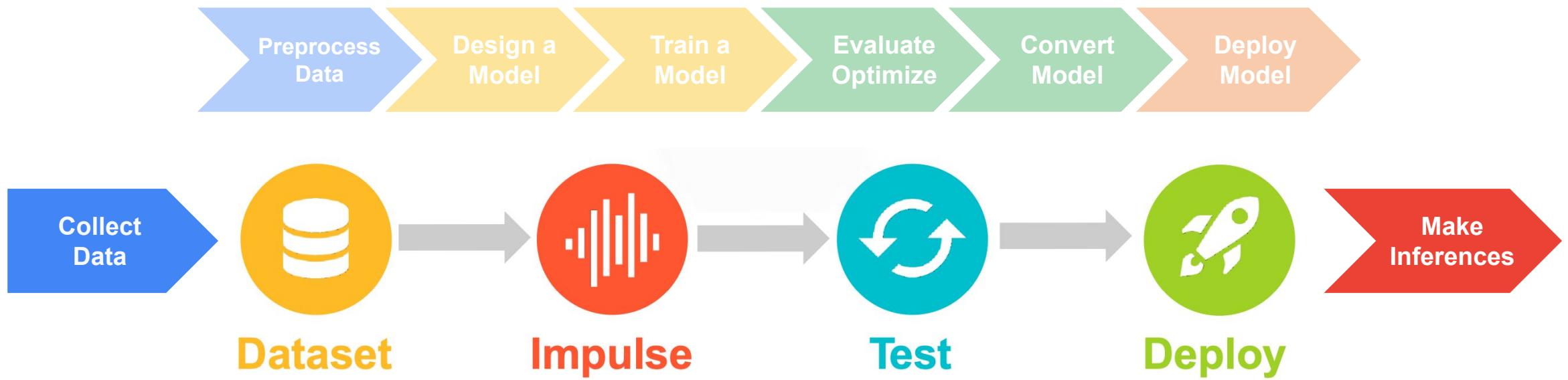


Machine Learning Workflow (“How”)

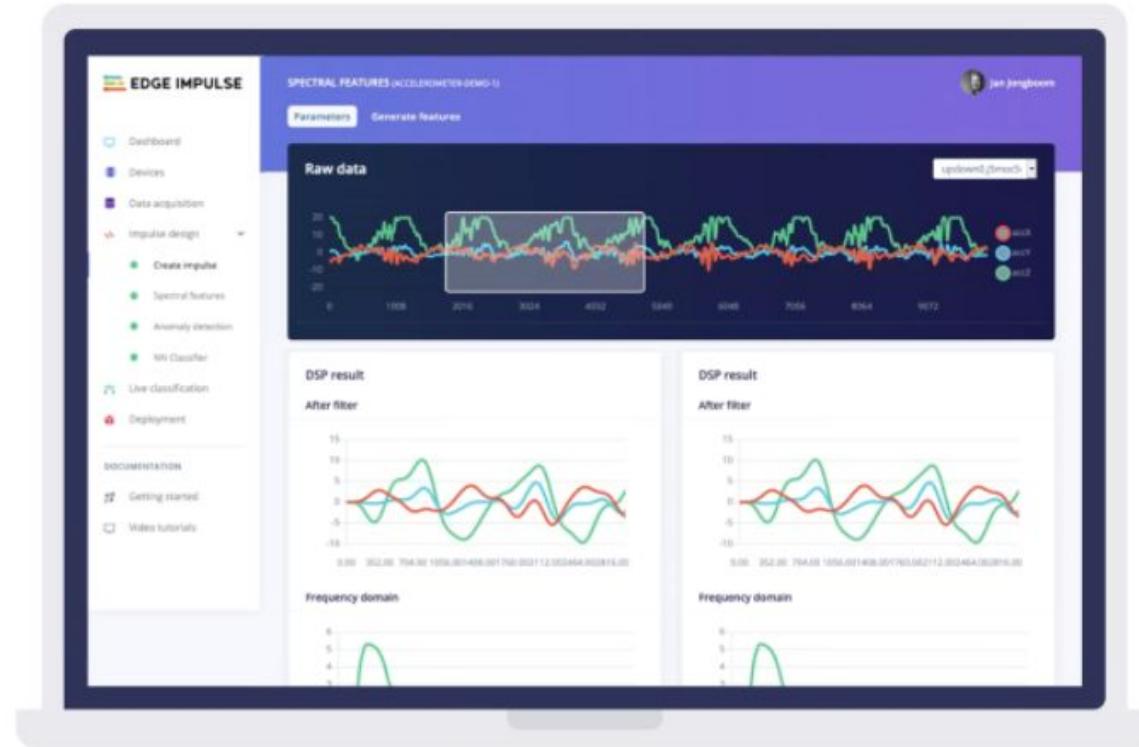
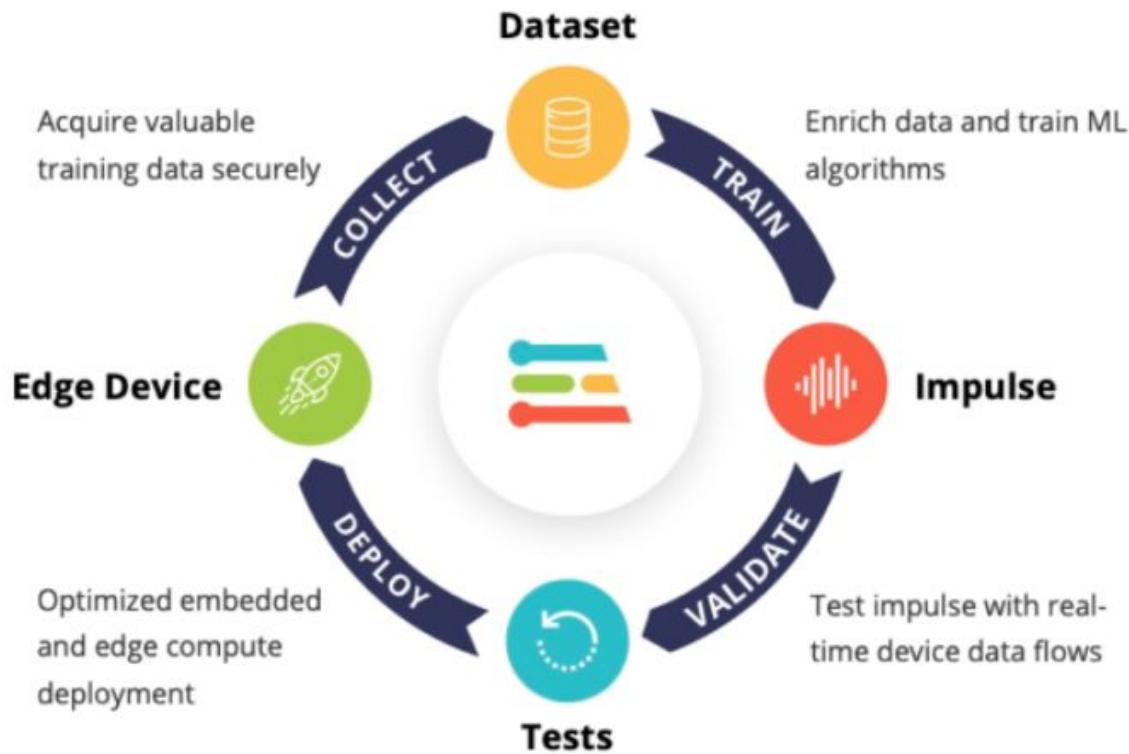


Machine Learning Workflow (“How”)





EI Studio - Embedded ML platform (“AutoML”)



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



Cifar10

Edge Impulse Studio



Dataset: <https://github.com/YoongiKim/CIFAR-10-images>

Join the biggest embedded ML event of the year, Imagine 2021. Register now for free! >

EDGE IMPULSE

Solutions Docs Forum Blog **Login** **Sign Up**

Making things smarter

Edge Impulse is the leading development platform for machine learning on edge devices, free for developers and trusted by enterprises.

Name

Email

Sign up for free

CONSERVATION
Elephant Activity

ELEPHANT
Human Proximity Confirmed

Trusted by thousands of embedded developers running critical machine learning projects across millions of data samples.

ML PROJECTS
▲ 39,151

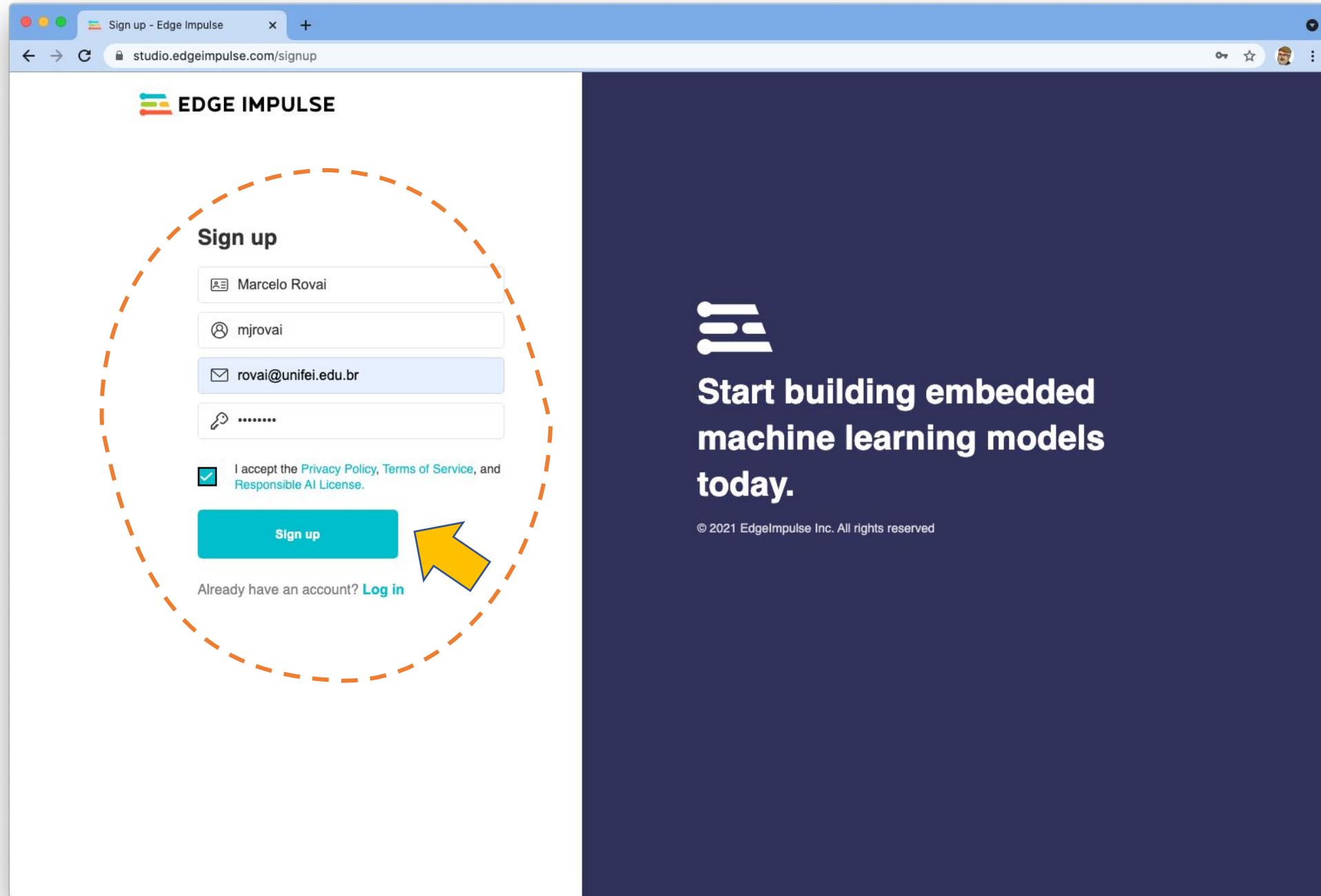
Jan 2020 Today

DATA SAMPLES
▲ 45.53 M

Jan 2020 Today

CLOUD JOBS
▲ 1.12 M

Jan 2020 Today



A screenshot of a web browser showing the Edge Impulse sign-up success page. The URL in the address bar is `studio.edgeimpulse.com/studio/signup-success`. The page features a white header with the Edge Impulse logo and a dark blue footer. The main content area on the left displays a success message and a button to build a model, while the right side has a call-to-action for building machine learning models.

Sign up successful!

Thanks Marcelo Rovai!

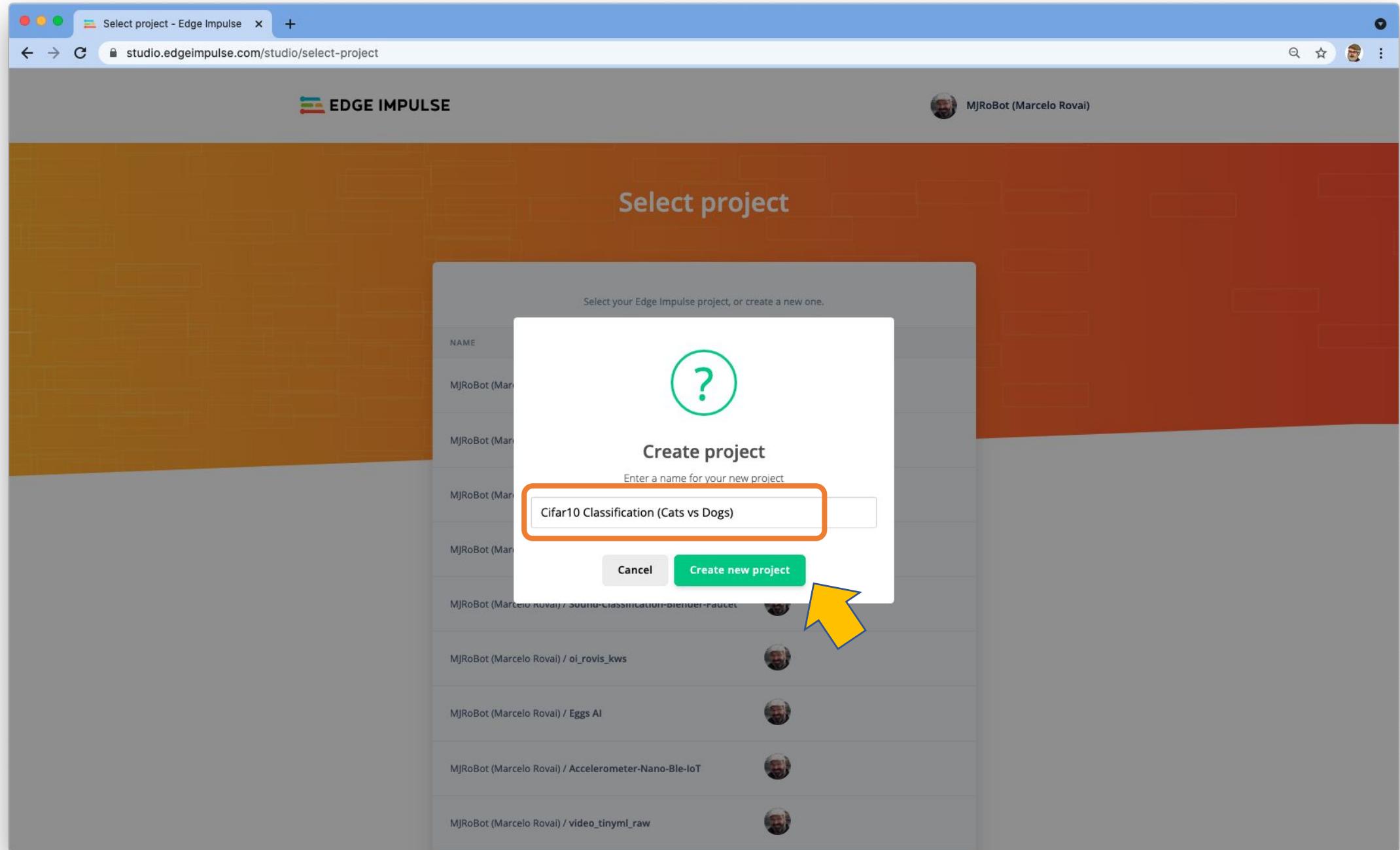
You have successfully signed up for Edge Impulse.

Click here to build your first ML model!

Re-send activation email

Start building embedded machine learning models today.

© 2021 EdgeImpulse Inc. All rights reserved



Select project - Edge Impulse

studio.edgeimpulse.com/studio/select-project

MJRoBot (Marcelo Rovai)

Select project

Select your Edge Impulse project, or create a new one.

NAME	COLLABORATORS
MJRoBot (Marcelo Rovai) / Cifar10 Classification (Cats vs Dogs)	
MJRoBot (Marcelo Rovai) / SoundClassification-Blender-Faucet	
MJRoBot (Marcelo Rovai) / ol_rovis_kws	
MJRoBot (Marcelo Rovai) / Eggs AI	
MJRoBot (Marcelo Rovai) / Accelerometer-Nano-Ble-IoT	
MJRoBot (Marcelo Rovai) / video_tinyml_raw	

Created project

Successfully created project "Cifar10 Classification (Cats vs Dogs)"

OK

A yellow arrow points from the bottom left towards the "OK" button on the modal dialog.

Dashboard - Cifar10 Classifica  +    

← → C  studio.edgeimpulse.com/studio/51076

EDGE IMPULSE

Project info Keys Export

MJRoBot (Marcelo Rovai)

Dashboard Devices Data acquisition Impulse design Create impulse EON Tuner Retrain model Live classification Model testing Versioning Deployment

GETTING STARTED Documentation Forums

MJRoBot (Marcelo Rovai) / Cifar10 Classification (Cats vs Dogs)

This is your Edge Impulse project.

Welcome to your new Edge Impulse project!

You're ready to add real intelligence to your edge devices. Let's set up your project. What type of data are you dealing with?

Creating your first impulse

Acquire data

Accelerometer data Analyze movement of your device in real-time to predict machine failure, detect human gestures, or monitor rotating machines.

Audio Listen to what's happening around you to create voice interfaces, listen to keywords, detect audible events, or to hear what's happening around your device.

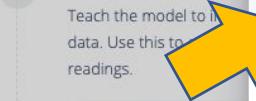
Images Add sight to your sensors with image classification or object detection - to detect humans and animals, monitor production lines or track objects.

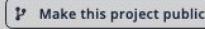
Something else Different sensor? No problem! You can collect and import data from any sensor, from environmental sensors to radars - and deploy your trained model back to virtually any device.

I know what I'm doing, hide this wizard!

Deploy

Package the complete impulse up, from signal processing code to trained model, and deploy it on your device. This ensures that the impulse runs with low latency and without requiring a network connection.

Sharing Your project is private. 

Summary DEVICES CONNECTED 0 DATA COLLECTED

Collaborators MJRoBot (Marcelo Rovai) OWNER

Project info

Dashboard - Cifar10 Classifica x +

studio.edgeimpulse.com/studio/51076

EDGE IMPULSE

Project info Keys Export

MJRoBot (Marcelo Rovai)

Dashboard Devices Data acquisition Impulse design Create impulse EON Tuner Retrain model Live classification Model testing Versioning Deployment

GETTING STARTED Documentation Forums

MJRoBot (Marcelo Rovai) / Cifar10 Classification (Cats vs Dogs)

This is your Edge Impulse project.

Welcome to your new Edge Impulse project!

Great! What do you want to detect?

Creating your first impulse

Acquire data

Every Machine Learning project needs data. You can collect it from a development board or your smartphone.

LET'S COLLECT SOME DATA

Design an impulse

Teach the model to interpret your data. Use this to categorize sensor readings.

GETTING STARTED: CONCEPT

GETTING STARTED: RESULTS

GETTING STARTED: ADVICE

I know what I'm doing, hide this wizard!

Deploy

Package the complete impulse up, from signal processing code to trained model, and deploy it on your device. This ensures that the impulse runs with low latency and without requiring a network connection.

DEPLOY YOUR MODEL

Sharing

Your project is private.

Make this project public

Summary

DEVICES CONNECTED 0

DATA COLLECTED -

Collaborators

MJRoBot (Marcelo Rovai) OWNER

Project info



18

Dashboard - Cifar10 Classifica +

studio.edgeimpulse.com/studio/51076

EDGE IMPULSE

Project info Keys Export

MJRoBot (Marcelo Rovai)

Dashboard Devices Data acquisition Impulse design Create impulse EON Tuner Retrain model Live classification Model testing Versioning Deployment

GETTING STARTED Documentation Forums

MJRoBot (Marcelo Rovai) / Cifar10 Classification (Cats vs Dogs)

This is your Edge Impulse project.

Welcome to your new Edge Impulse project!

Great! Here's how you can get started with image classification:

Connect a development board

Get started with real hardware from Arduino, OpenMV, Eta Compute, Himax, Raspberry Pi or NVIDIA to quickly build a custom image dataset.

[Connect your development board](#)

Acquire data

Every Machine Learning project needs data. You can collect it using a development board or your smartphone.

[LET'S COLLECT SOME DATA](#)

Import existing data

If you already have images in JPG or PNG file format, you can upload it to Edge Impulse through the web interface or using the Edge Impulse CLI.

[Go to the uploader](#)

Design an impulse

Teach the model to interpret your sensor data. Use this to categorize your sensor readings.

[GETTING STARTED: CONCEPTS](#) [GETTING STARTED: RESULTS](#) [GETTING STARTED: ADVICE](#)

Tutorial: adding sight to your sensors

Follow our end-to-end tutorial to collect data, train a model, and deploy it back to your device to analyze images in realtime.

[Read the tutorial](#)

I know what I'm doing, hide this wizard!

[Let's get started!](#)

Deploy

Package the complete impulse up, from signal processing code to trained model, and deploy it on your device. This ensures that the impulse runs with low latency and without requiring a network connection.

[DEPLOY YOUR MODEL](#)

Sharing

Your project is private.

[Make this project public](#)

Summary

DEVICES CONNECTED 0

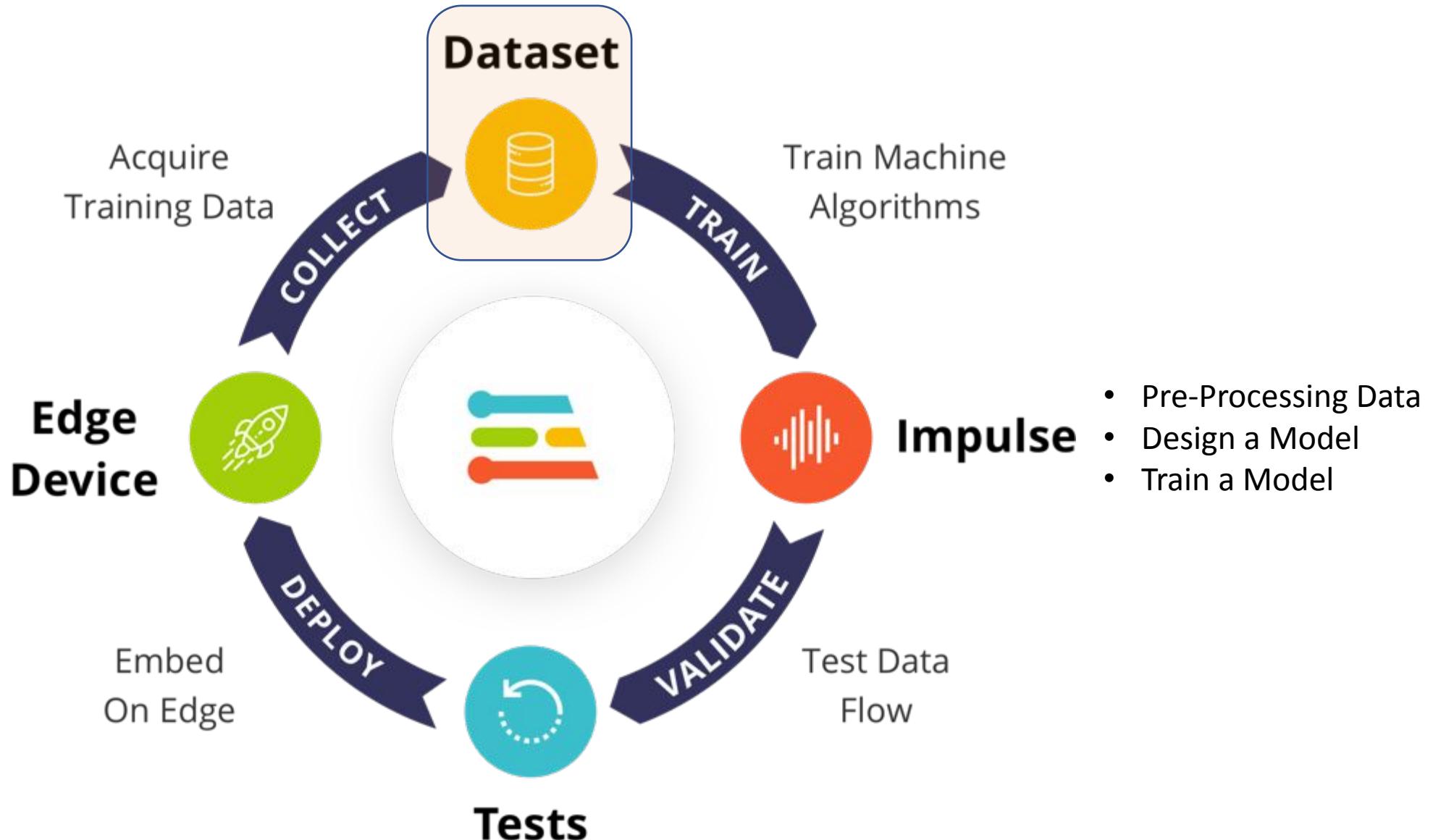
DATA COLLECTED -

Collaborators

MJRoBot (Marcelo Rovai) OWNER

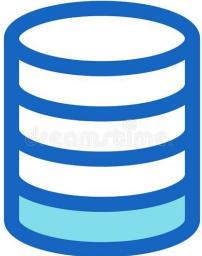
Project info

https://studio.edgeimpulse.com/studio/51076/upload

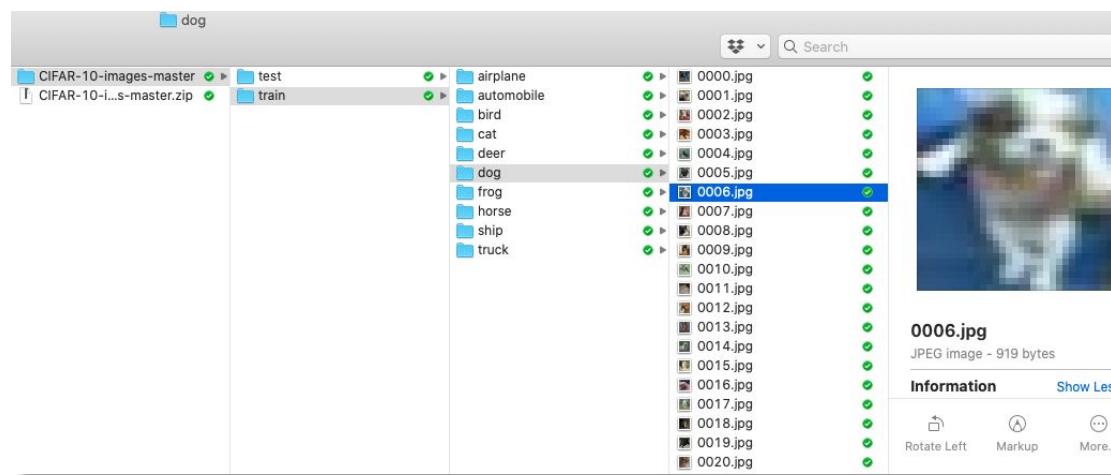


Download Dataset

<https://github.com/YoongiKim/CIFAR-10-images>



The screenshot shows a GitHub repository page for 'CIFAR-10-images'. At the top, there are buttons for 'Code' (highlighted with an orange box), 'Go to file', 'Add file', and 'About'. Below the buttons, there's a 'Clone' section with 'HTTPS' selected, showing the URL 'https://github.com/YoongiKim/CIFAR-10-images'. There are also links for 'Open with GitHub Desktop' and 'Download ZIP'. A large yellow arrow points from the monitor icon at the bottom to this 'Download ZIP' button. The main content area displays the repository's README.md file, which contains the text 'CIFAR-10 raw jpeg images' and 'You can just clone this repository to use dataset.'



Collect Data

Upload data - Cifar10_Image_C

stUDIO.edgeimpulse.com/studio/51070/upload

EDGE IMPULSE

Dashboard

Devices

Data acquisition

Impulse design

- Create impulse
- Image
- NN Classifier

EON Tuner

Retrain model

Live classification

Model testing

Versioning

Deployment

GETTING STARTED

Documentation

Forums

UPLOAD DATA (CIFAR10_IMAGE_CLASSIFICATION)

Upload existing data

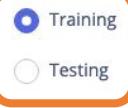
You can upload existing data to your project in the [Data Acquisition Format](#) (CBOR, JSON, CSV), or as WAV, JPG or PNG files.

Select files 

No file chosen

Upload into category

Automatically split between training and testing ?

Training 

Testing

Label

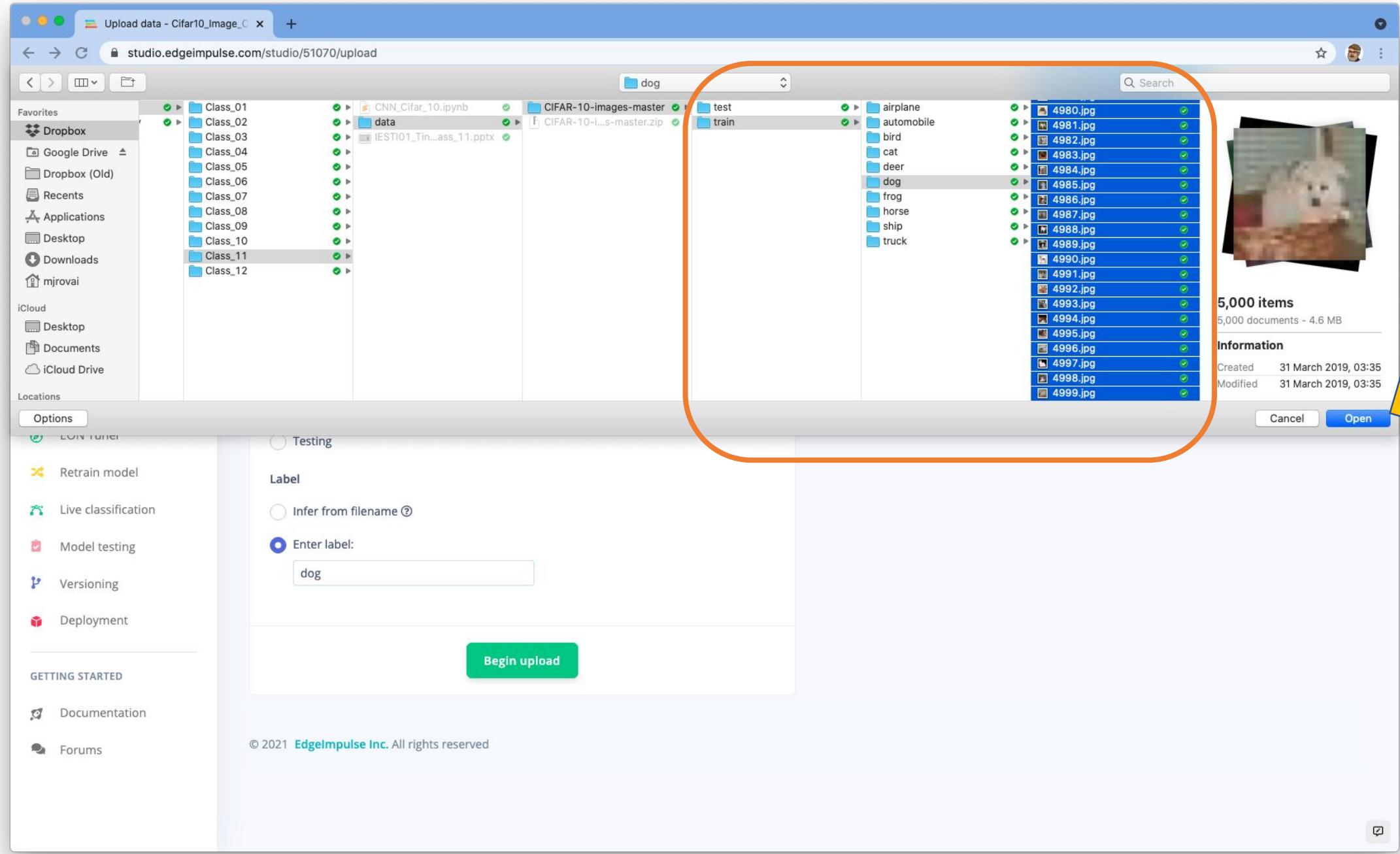
Infer from filename ?

Enter label: 

dog

Begin upload

MJRoBot (Marcelo Rovai)



Collect Data

Upload data - Cifar10_Image_C

stUDIO.edgeimpulse.com/studio/51070/upload

EDGE IMPULSE

MJRoBot (Marcelo Rovai)

Dashboard

Devices

Data acquisition

Impulse design

- Create impulse
- Image
- NN Classifier

EON Tuner

Retrain model

Live classification

Model testing

Versioning

Deployment

GETTING STARTED

Documentation

Forums

UPLOAD DATA (CIFAR10_IMAGE_CLASSIFICATION)

Upload existing data

You can upload existing data to your project in the [Data Acquisition Format](#) (CBOR, JSON, CSV), or as WAV, JPG or PNG files.

Select files

5000 files

Upload into category

Automatically split between training and testing ?

Training

Testing

Label

Infer from filename ?

Enter label:

dog

Begin upload

Collect Data

EDGE IMPULSE

Data acquisition - Cifar10_Img

Training data Test data

Did you know? You can capture data from any device or development board, or upload your existing datasets - Show options

DATA COLLECTED
10,000 items

TRAIN / TEST SPLIT
83% / 17%

Collected data

SAMPLE NAME	LABEL	ADDED	LENGTH
4999.jpg.2hfb34u0	dog	Today, 11:14:02	-
4997.jpg.2hfb34n1	dog	Today, 11:14:01	-
4991.jpg.2hfb34mm	dog	Today, 11:14:01	-
4993.jpg.2hfb34n8	dog	Today, 11:14:01	-
4996.jpg.2hfb34o3	dog	Today, 11:14:01	-
4987.jpg.2hfb34ni	dog	Today, 11:14:01	-
4995.jpg.2hfb34nq	dog	Today, 11:14:01	-
4988.jpg.2hfb34mh	dog	Today, 11:14:01	-
4990.jpg.2hfb34n8	dog	Today, 11:14:01	-
4989.jpg.2hfb34n8	dog	Today, 11:14:01	-
4992.jpg.2hfb34nq	dog	Today, 11:14:01	-
4994.jpg.2hfb34mr	dog	Today, 11:14:01	-

Record new data Connect using WebUSB

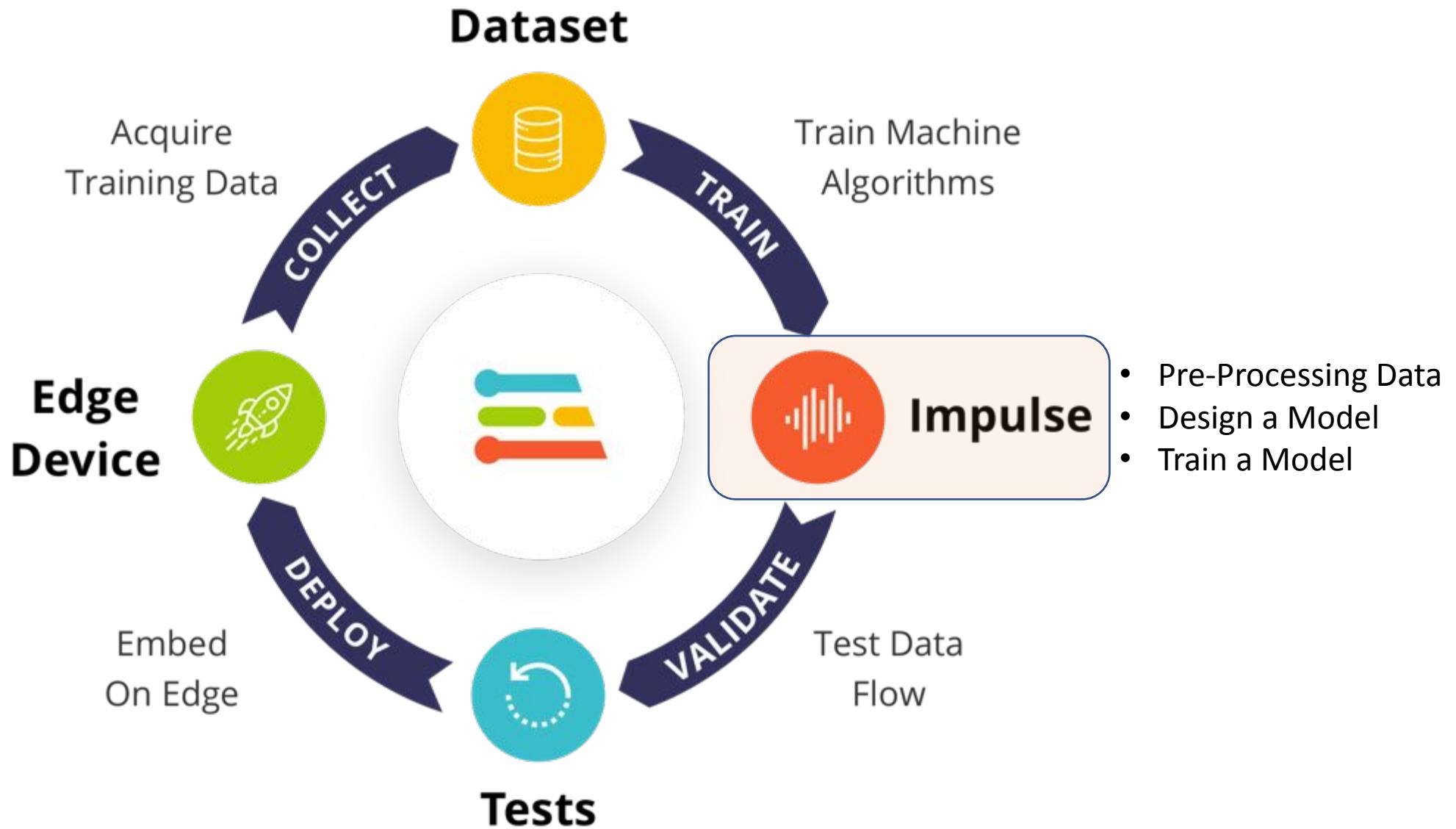
No devices connected to the remote management API.

RAW DATA
4999.jpg.2hfb34u0



834

25



Create impulse - Cifar10_Image

studio.edgeimpulse.com/studio/51070/create-impulse

EDGE IMPULSE

CREATE IMPULSE (CIFAR10_IMAGE_CLASSIFICATION)

An impulse takes raw data, uses signal processing to extract features, and then uses a learning block to classify new data.

Image data

Axes

Image width: 32 Image height: 32

Resize mode: Fit longest axis

For optimal accuracy with transfer learning blocks, use a 96x96 or 160x160 image size.

Image

Name: Image

Input axes: image

Classification (Keras)

Name: NN Classifier

Input features: Image

Output features: 2 (cat, dog)

Output features

2 (cat, dog)

Save impulse

Image width and height fields highlighted with a blue border.

Yellow arrow pointing to the "Save impulse" button.

Add a processing block

Add a learning block

© 2021 EdgeImpulse Inc. All rights reserved

Preprocess Data

Image - Cifar10_Image_Classifi x +

studio.edgeimpulse.com/studio/51070/dsp/image/3

EDGE IMPULSE

IMAGE (CIFAR10_IMAGE_CLASSIFICATION)

#1 Click to set a description for this version

Parameters Generate features

Raw data

4999.jpg.2hfb34u0 (dog)

Raw features

0xa9aaa4, 0xafb0aa, 0xabaca6, 0xa9aaa4, 0xaebl8, 0xabaee5, 0xa6a9a0, 0xabaee5, 0xabaee3, 0xaaad...

Parameters

Image

Color depth

RGB

Save parameters

DSP result

Image

Processed features

0.6627, 0.6667, 0.6431, 0.6863, 0.6902, 0.6667, 0.6706, 0.6745, 0.6510, 0.6627, 0.6667, 0.6431, ...

On-device performance

PROCESSING TIME 9 ms.

PEAK RAM USAGE 4 KB

© 2021 EdgeImpulse Inc. All rights reserved

Preprocess Data

Image - Cifar10_Image_Classifi x +

studio.edgeimpulse.com/studio/51070/dsp/image/3/generate-features

EDGE IMPULSE

IMAGE (CIFAR10_IMAGE_CLASSIFICATION)
#1 Click to set a description for this version

Parameters Generate features

Training set

Data in training set 10,000 items

Classes 2 (cat, dog)

Generate features

Feature explorer (10,000 samples) ②

X Axis Y Axis Z Axis

Visualization layer 1 Visualization layer 2 Visualization layer 3

cat dog

8.5
8
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0

Visualization layer 3
Visualization layer 2
Visualization layer 1

On-device performance ②

PROCESSING TIME 9 ms.

PEAK RAM USAGE 4 KB

© 2021 EdgeImpulse Inc. All rights reserved

EDGE IMPULSE

NN CLASSIFIER (CIFAR10_IMAGE_CLASSIFICATION)
#2 ▾ Model - Same as done in class (IESTI01) ★ Primary version

Dashboard
Devices
Data acquisition
Impulse design
Create impulse
Image
NN Classifier
EON Tuner
Retrain model
Live classification
Design a Model
Versioning
Deployment
Documentation
Forums

Neural Network settings

Training settings

Number of training cycles ① Learning rate ②

Neural network architecture

Input layer (3,072 features)

2D conv / pool layer (32 filters, 3 kernel size, 1 layer)

2D conv / pool layer (64 filters, 3 kernel size, 1 layer)

Flatten layer

Dense layer (64 neurons)

Add an extra layer

Output layer (2 features)

Start training

Training output

Model version: ② Quantized (int8)

Model

Last training performance (validation set)

ACCURACY 74.4% LOSS 0.56

Confusion matrix (validation set)

	CAT	DOG
CAT	80.6%	19.4%
DOG	32.1%	67.9%
F1 SCORE	0.76	0.72

Feature explorer (full training set) ③

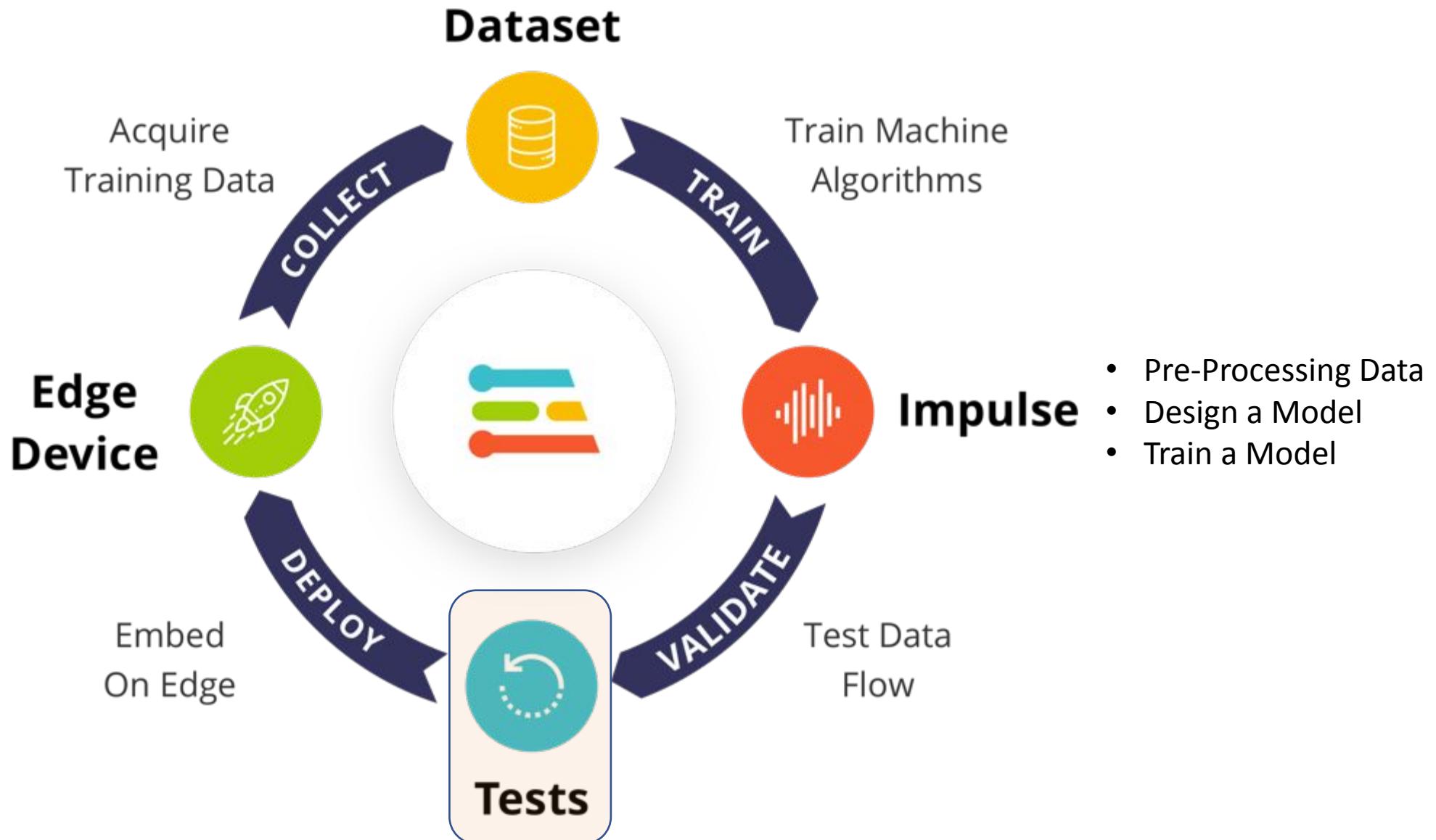
cat - correct
dog - correct
cat - incorrect
dog - incorrect

Visualization layer 3
Visualization layer 2
Visualization layer 1

On-device performance ④

INFERENCING TIME 838 ms. PEAK RAM USAGE 44.7K FLASH USAGE 308.2K

30



Model testing - Cifar10_Image... Upload data - Cifar10 Classific... | +

studio.edgeimpulse.com/studio/51070/validation

EDGE IMPULSE

- Dashboard
- Devices
- Data acquisition
- Impulse design
 - Create impulse
 - Image
 - NN Classifier
- EON Tuner
- Retrain model
- Live classification
 - Model testing
 - Versioning
- Deployment

Evaluate Optimize

Test data

Set the 'expected outcome' for each sample to the desired outcome to automatically score the impulse.

SAMPLE NAME	EXPECTED OUTCOME	LENGTH	ACCURACY	RESULT	⋮
testing.2hfe5uat	testing	-		1 cat	⋮
0999.jpg.2hfb7a...	cat	-	100%	1 cat	⋮
0997.jpg.2hfb7a...	cat	-	100%	1 cat	⋮
0998.jpg.2hfb7a...	cat	-	100%	1 cat	⋮
0996.jpg.2hfb7a...	cat	-	0%	1 uncertain	⋮
0993.jpg.2hfb7a...	cat	-	100%	1 cat	⋮
0995.jpg.2hfb7a...	cat	-	0%	1 dog	⋮
0991.jpg.2hfb7a...	cat	-	100%	1 cat	⋮
0994.jpg.2hfb7a...	cat	-	100%	1 cat	⋮
0992.jpg.2hfb7a...	cat	-	100%	1 cat	⋮
0990.jpg.2hfb79...	cat	-	100%	1 cat	⋮
0988.jpg.2hfb79...	cat	-	100%	1 cat	⋮
0985.jpg.2hfb79...	cat	-	0%	1 uncertain	⋮

Model testing output

Classifying data for NN classifier...
 Copying features from processing blocks...
 Copying features from DSP block...
 Copying features from DSP block OK
 Copying features from processing blocks OK

Classifying data for float32 model...
 Scheduling job in cluster...
 Job started
 Classifying data for NN Classifier OK

Job completed

Model testing results

ACCURACY **64.95%** %

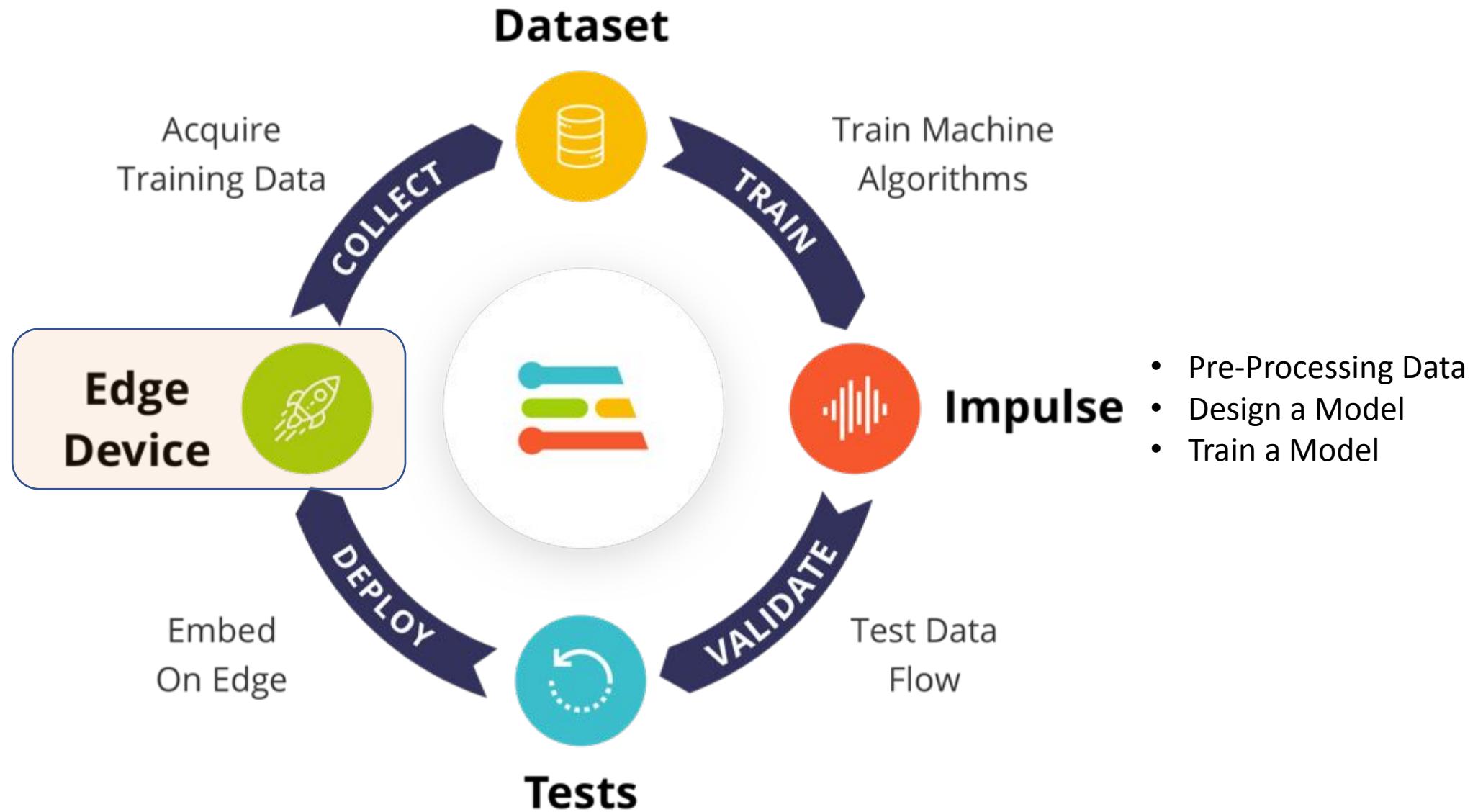
	CAT	DOG	UNCERTAIN
CAT	67.7%	16%	16.3%
DOG	21.2%	62.2%	16.6%
F1 SCORE	0.72	0.70	

Feature explorer ⓘ

Legend:

- cat - correct
- dog - correct
- cat - incorrect
- dog - incorrect
- testing

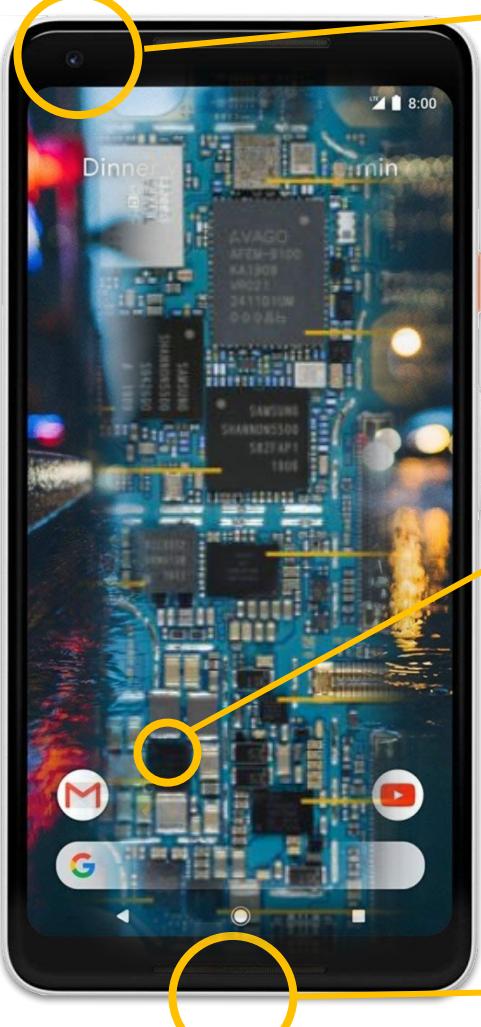
The plot shows a complex, overlapping distribution of points in a 3D space defined by 'Visualization layer 1', 'Visualization layer 2', and 'Visualization layer 3'. Points are colored according to the legend: green for correct classifications and red for incorrect ones. A significant number of points fall into the 'uncertain' category, appearing as a dense cluster near the origin.



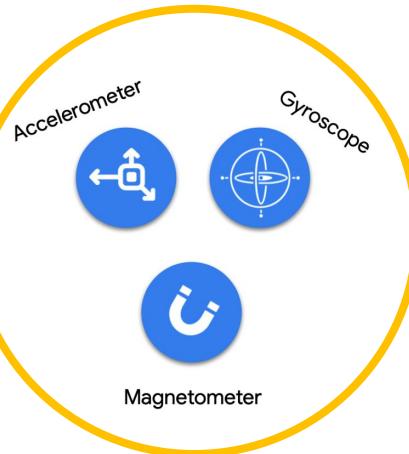
Edge Device



& Sensors



Camera



Accelerometer



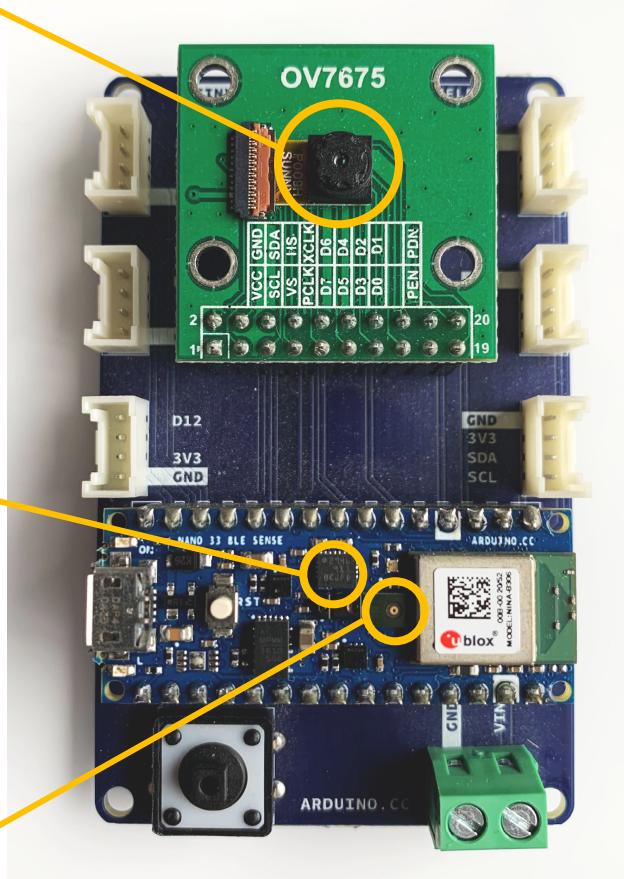
Gyroscope



Magnetometer



Microphone



The screenshot shows the Edge Impulse web interface with a modal window titled "Collect data". The modal provides instructions on how to collect data, listing several methods:

- Connect a fully supported development board**: Get started with real hardware from a wide range of silicon vendors - fully supported by Edge Impulse. Includes a "Browse dev boards" button.
- Use your mobile phone**: Use your mobile phone to capture movement, audio or images, and even run your trained model locally. No app required. Includes a "Show QR code" button highlighted with a red dashed box and an orange arrow pointing to it.
- Use your computer**: Capture audio or images from your webcam or microphone, or from an external audio device. Includes a "Collect data" button.
- Data from any device with the data forwarder**: Capture data from any device or development board over a serial connection, in 10 lines of code. Includes a "Show docs" button.
- Upload data**: Already have data? You can upload your existing datasets directly in WAV, JPG, PNG, CBOR, CSV or JSON format. Includes a "Go to the uploader" button.
- Integrate with your cloud**: The enterprise version of Edge Impulse integrates directly with the data stored in your cloud platform. Includes a "Contact us" button.

The left sidebar shows the navigation menu with "Devices" selected. The top right corner shows the user profile of Marcelo Rovai.

Devices - TinyML4D - Project

studio.edgeimpulse.com/studio/49268/devices

EDGE IMPULSE

DEVICES (TINYML4D - PROJECT SETUP)

Your devices

+ Connect a new device

These are devices that are connected to the Edge Impulse remote management API, or have posted data to the ingestion SDK.

Collect data

You can collect data from any smartphone. From your smartphone go to [this URL](#), or scan the QR code below.



© 2021 Ed

Devices

Dashboard

Data acquisition

Impulse design

Create impulse

Retrain model

Live classification

Model testing

Versioning

Deployment

GETTING STARTED

Documentation

Forums

Marcelo Rovai

WEBSITE QR CODE
Open "edgeimpulse.com" in Safari



Devices - TinyML4D - Project

studio.edgeimpulse.com/studio/49268/devices

EDGE IMPULSE

Devices

Your devices

These are devices that are connected to the Edge Impulse remote management API, or have posted data to the ingestion SDK.

NAME	ID	TYPE	SENSORS	REMO...	LAST SEEN
phone_kq6ray4k	phone_kq6ray4k	MOBILE CLIENT	Accelerometer, Microphone	...	Today, 12:06:04

+ Connect a new device

Collect data

Device phone_kq6ray4k is now connected

Get started!

© 2021 Edge Impulse Inc.

Camera 12:07 22% smartphone.edgeimpulse.com

Data collection

Connected as phone_kq6ray4k

You can collect data from this

A large yellow arrow points upwards towards the 'Get started!' button in the 'Collect data' modal.

Devices - TinyML4D - Project

studio.edgeimpulse.com/studio/49268/devices

EDGE IMPULSE

DEVICES (TINYML4D - PROJECT SETUP)

Marcelo Rovai

Your devices

+ Connect a new device

These are devices that are connected to the Edge Impulse remote management API, or have posted data to the ingestion SDK.

NAME	ID	TYPE	SENSORS	REMO...	LAST SEEN
phone_kq6ray4k	phone_kq6ray4k	MOBILE_CLIENT	Accelerometer, Microph...	●	Today, 12:06:04

© 2021 EdgeImpulse Inc. All rights reserved

Dashboard

Devices (highlighted with orange dashed box)

Data acquisition

Impulse design

- Create impulse

Retrain model

Live classification

Model testing

Versioning

Deployment

GETTING STARTED

Documentation

Forums

Camera 12:07 22%

smartphone.edgeimpulse.com

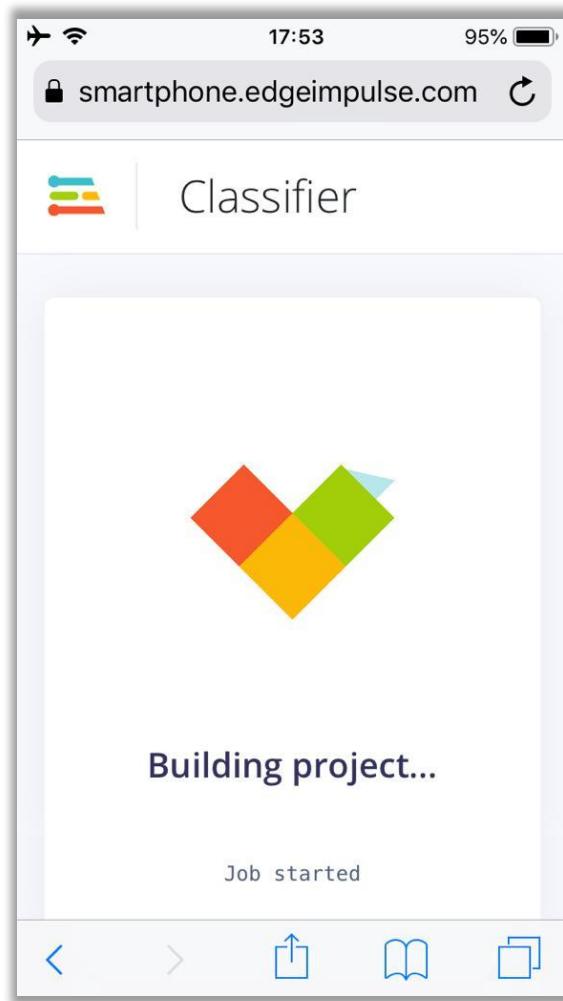
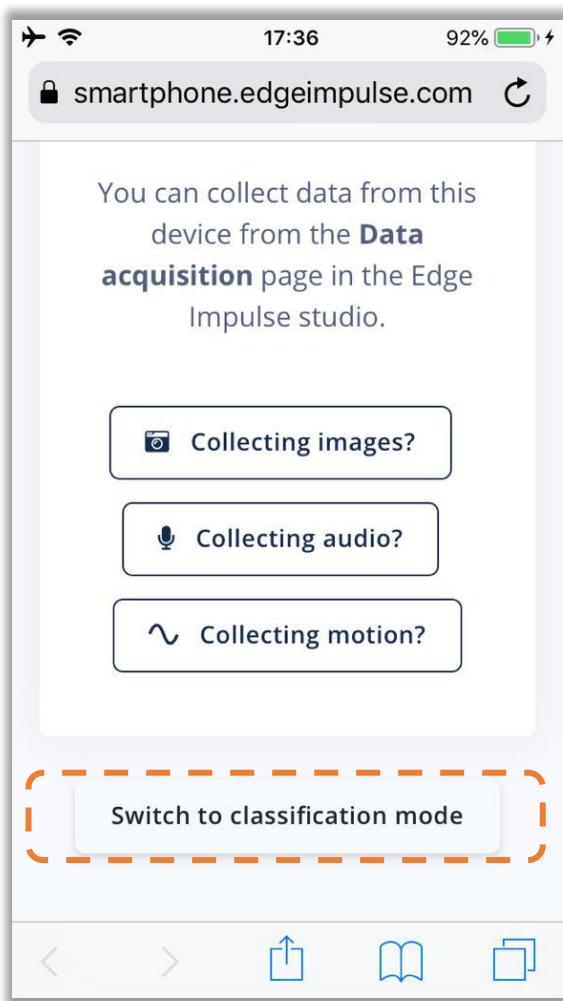
Data collection

Connected as phone_kq6ray4k

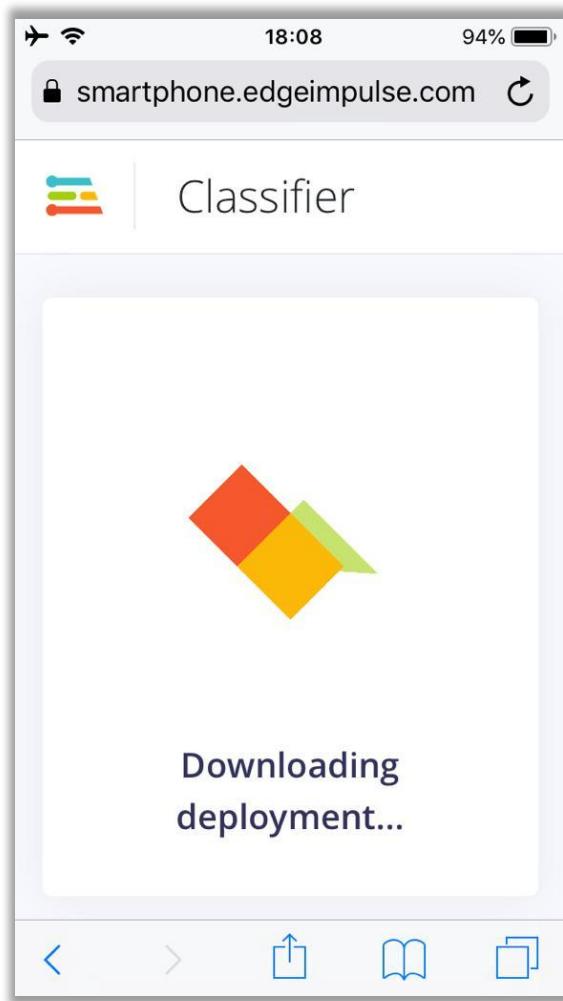
You can collect data from this



Convert Model



Deploy Model



Make Inferences

The figure displays four sequential screenshots of a mobile application interface, illustrating the process of making inferences on images using a classifier. The interface includes a header bar with the URL ne.edgeimpulse.com, a camera icon, signal strength, battery level, and a refresh button.

Screenshot 1 (11:58): Shows a Beagle dog running. The classifier results table shows:

	CAT	DOG
16	0.03	0.97
15	0.49	0.51

Screenshot 2 (11:59): Shows a white dog with brown spots. The classifier results table shows:

	CAT	DOG
17	0.00	1.00
16	0.03	0.97

Screenshot 3 (12:04): Shows a fluffy light-colored cat. The classifier results table shows:

	CAT	DOG
39	0.94	0.06
38	0.98	0.02

Screenshot 4 (12:04): Shows a woman with a cat head. The classifier results table shows:

	CAT	DOG
40	0.64	0.36
39	0.94	0.06

Each screenshot includes a "Next photo" button at the bottom center and standard iOS navigation icons (back, forward, share, etc.) at the bottom.

EI Studio “under the hood” Code Time!

ei-cifar10_image_classification-nn-classifier-explained.ipynb



EDGE IMPULSE

NN CLASSIFIER (CIFAR10_IMAGE_CLASSIFICATION)

#2 ▾ Model - Same as done in class (IESTI01) ★ Primary version

Neural Network settings

Training settings

Neural network architecture

```

1 import tensorflow as tf
2 from tensorflow.keras.models import Sequential
3 from tensorflow.keras.layers import Dense, InputLayer, Dropout, Conv2D, Flatten,
4     Reshape, MaxPooling1D, MaxPooling2D, BatchNormalization, TimeDistributed
5
6 # model architecture
7 model = Sequential()
8 model.add(Conv2D(32, kernel_size=3, activation='relu', kernel_constraint=tf.keras
    .constraints.MaxNorm(1), padding='same'))
9 model.add(MaxPooling2D(pool_size=2, strides=2, padding='same'))
10 model.add(Conv2D(64, kernel_size=3, activation='relu', kernel_constraint=tf.keras
    .constraints.MaxNorm(1), padding='same'))
11 model.add(MaxPooling2D(pool_size=2, strides=2, padding='same'))
12 model.add(Flatten())
13 model.add(Dense(64, activation='relu',
    activity_regularizer=tf.keras.regularizers.l1(0.0001)))
14 model.add(Dense(classes, activation='softmax', name='y_pred'))
15
16 # this controls the learning rate
17 opt = Adam(lr=0.0005, beta_1=0.9, beta_2=0.999)
18 # this controls the batch size, or you can manipulate the tf.data.Dataset objects
19 # yourself
20 BATCH_SIZE = 32
21 train_dataset = train_dataset.batch(BATCH_SIZE, drop_remainder=False)
22 validation_dataset = validation_dataset.batch(BATCH_SIZE, drop_remainder=False)
23 callbacks.append(BatchLoggerCallback(BATCH_SIZE, train_sample_count))
24
25 # train the neural network
26 model.compile(loss='categorical_crossentropy', optimizer=opt, metrics=['accuracy'])

```

Start training

Training output

Model

Model version: ② Quantized (int8)

Last training performance (validation set)

ACCURACY 74.4%

LOSS 0.56

Confusion matrix (validation set)

	CAT	DOG
CAT	80.6%	19.4%
DOG	32.1%	67.9%
F1 SCORE	0.76	0.72

Feature explorer (full training set) ②

- cat - correct
- dog - correct
- cat - incorrect
- dog - incorrect

On-device performance ②

INFERENCING TIME 1,048 ms.

PEAK RAM USAGE 44.7K

FLASH USAGE 308.2K

NN Classifier - Cifar10_Image... X Upload data - Cifar10 Classific... X computer icon - Pesquisa Go... X +

studio.edgeimpulse.com/studio/51070/learning/keras/6

EDGE IMPULSE

NN CLASSIFIER (CIFAR10_IMAGE_CLASSIFICATION)
#2 ▾ Model - Same as done in class (IESTI01) ★ Primary version

MJRoBot (Marcelo Rovai)

Neural Network settings

Training settings

Number of training cycles ②

Learning rate ②

Switch to Keras (expert mode)

Edit as iPython notebook

Training output

Model

Last training performance (validation set)

ACCURACY 74.4%

LOSS 0.56

Confusion matrix (validation set)

	CAT	DOG
CAT	80.6%	19.4%
DOG	32.1%	67.9%
F1 SCORE	0.76	0.72

Feature explorer (full training set) ②

cat - correct
dog - correct
cat - incorrect
dog - incorrect

Visualization layer 3

Visualization layer 2

On-device performance ②

INFERENCING TIME 1,048 ms.

PEAK RAM USAGE 44.7K

FLASH USAGE 308.2K

https://studio.edgeimpulse.com/v1/api/51070/training/keras/6/ipynb

Switch to Keras (expert mode)

Edit as iPython notebook

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 especially 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!



UNIFEI