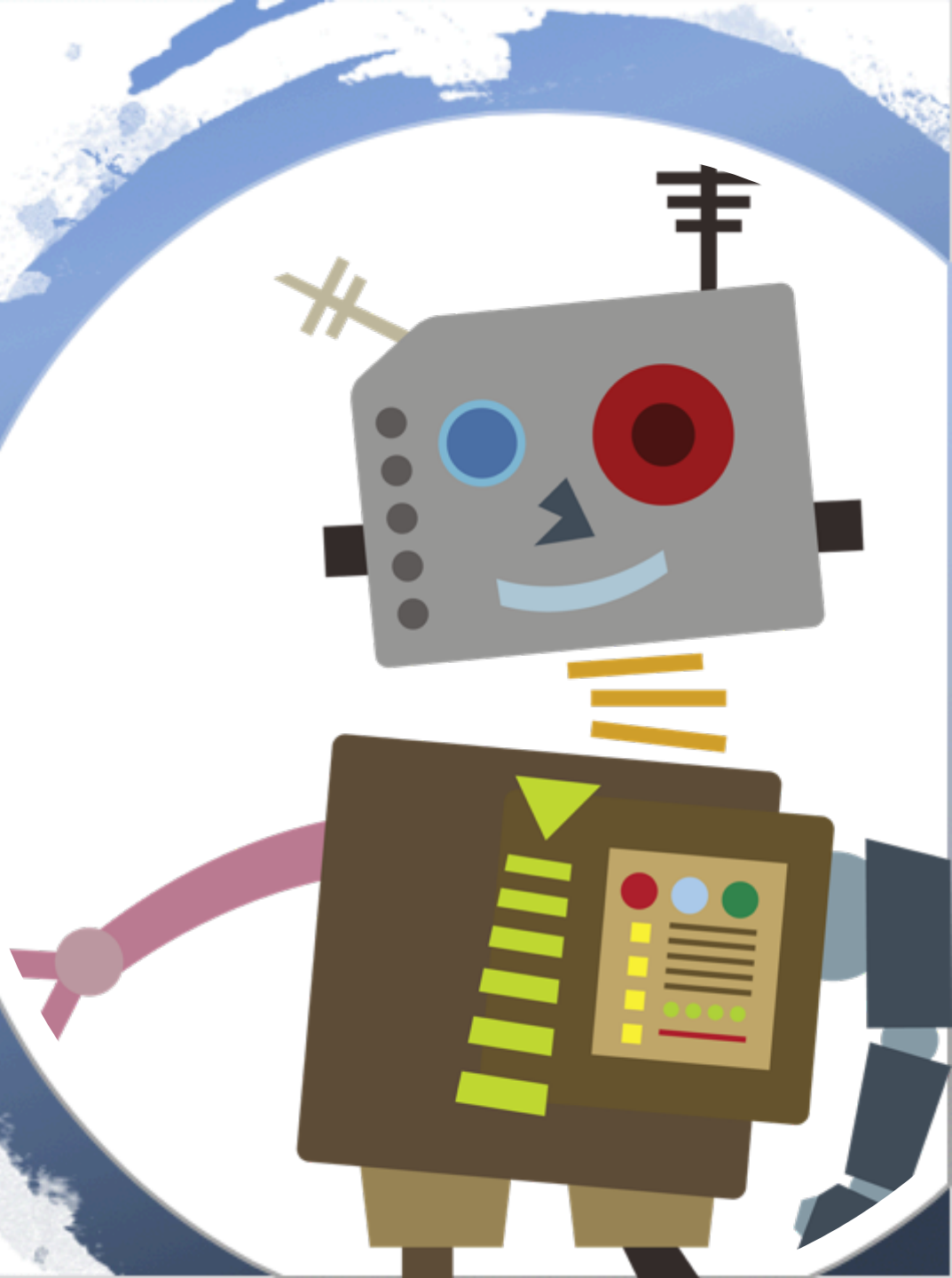
Exploring AI at the Edge!



By Marcelo Rovai

Meetup

IoT hacking & Data Science Innovators

August 27th, 2020

Marcelo Rovai

Brazilian from São Paulo, Master in Data Science by UDD, Chile, and MBA by IBMEC (INSPER), Brazil. Marcelo graduated in 1982 as an Engineer from UNIFEI with pos graduation by Poli/USP, both in Brazil. He worked as a teacher, engineer, and executive in several companies in the technology area such as AVIBRAS Aeroespacial, SID Informática, ATT-GIS, NCR, DELL, COMPAQ (HP) and more recently at IGT where he continues as Senior Advisor.

In 2016, Marcelo began writing about electronics, publishing his works in sites of the area as MJRoBot.org (Editor/Writer), Hackster.io (#1 Contributor), Instructables.com, and Medium.com (TDS – Towards Data Science). Besides winning several Instructables competitions in the areas of electronics, robotics, and IoT.

Marcelo lives with his wife Ilza in Santiago, Chile, where he divides his time between his consultant work and sharing ideas in the field of Data Science, Electronics, IoT, Physical Computing and Robotics.

Data → … → Value

Meetup

IoT hacking & Data Science Innovators

Data → ...→ Value

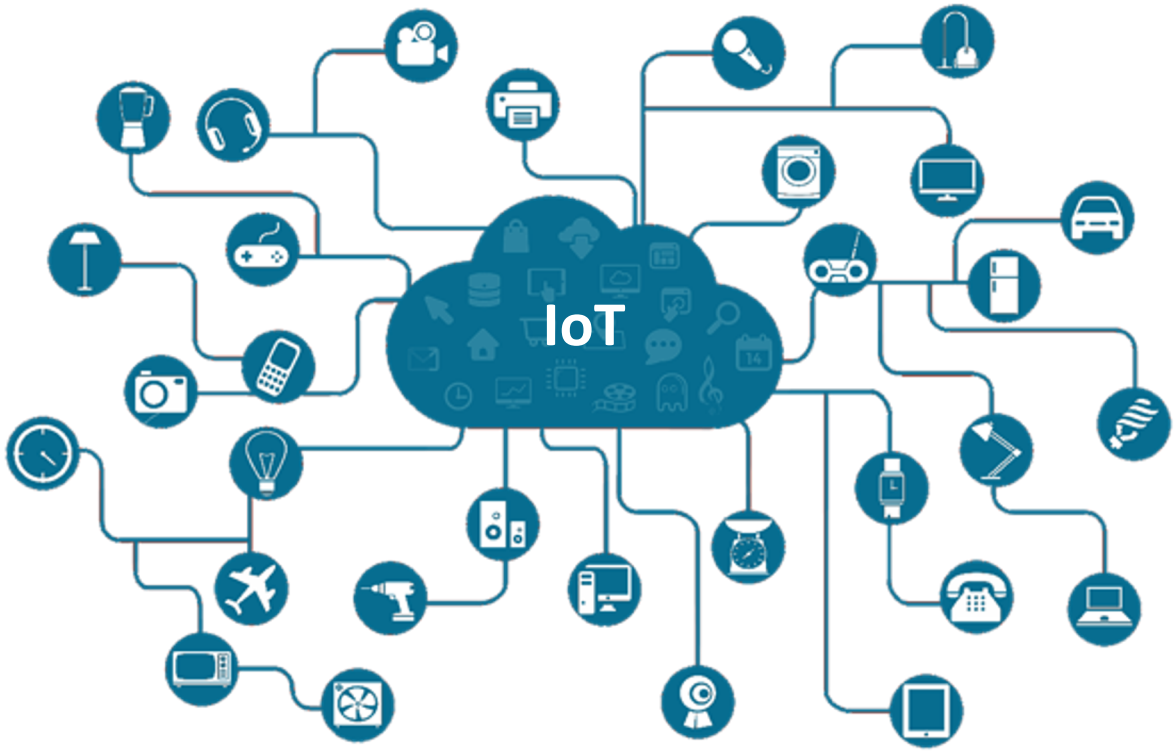
Meetup

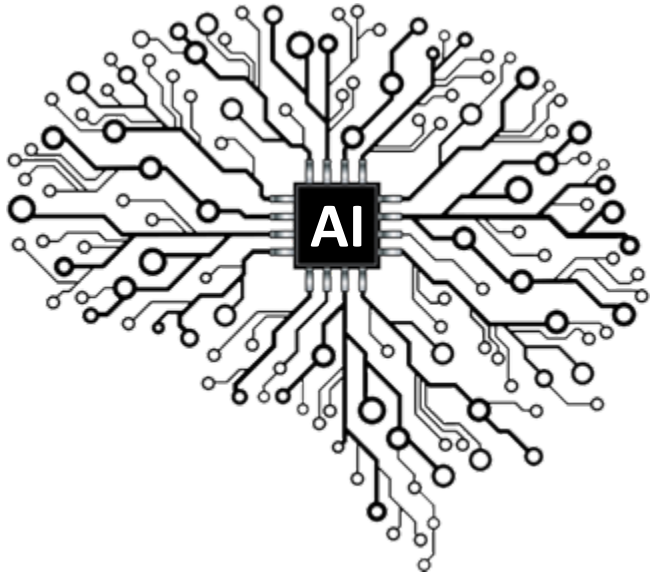
IoT hacking & Data Science Innovators

Data (IoT) ➔ Value (AI)

Meetup

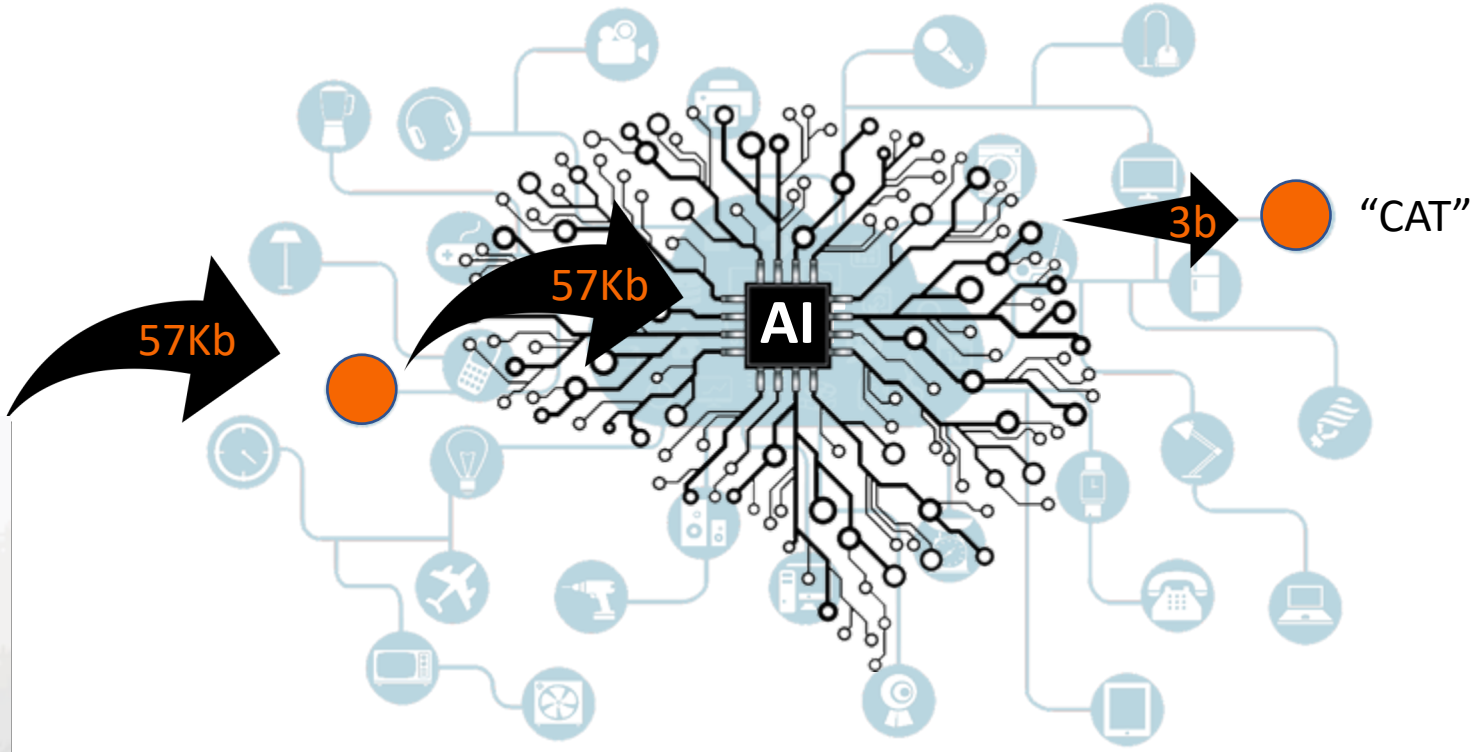
IoT hacking & Data Science Innovators





Data (IoT) ➔ Value (AI)

Meetup
IoT hacking & Data Science Innovators

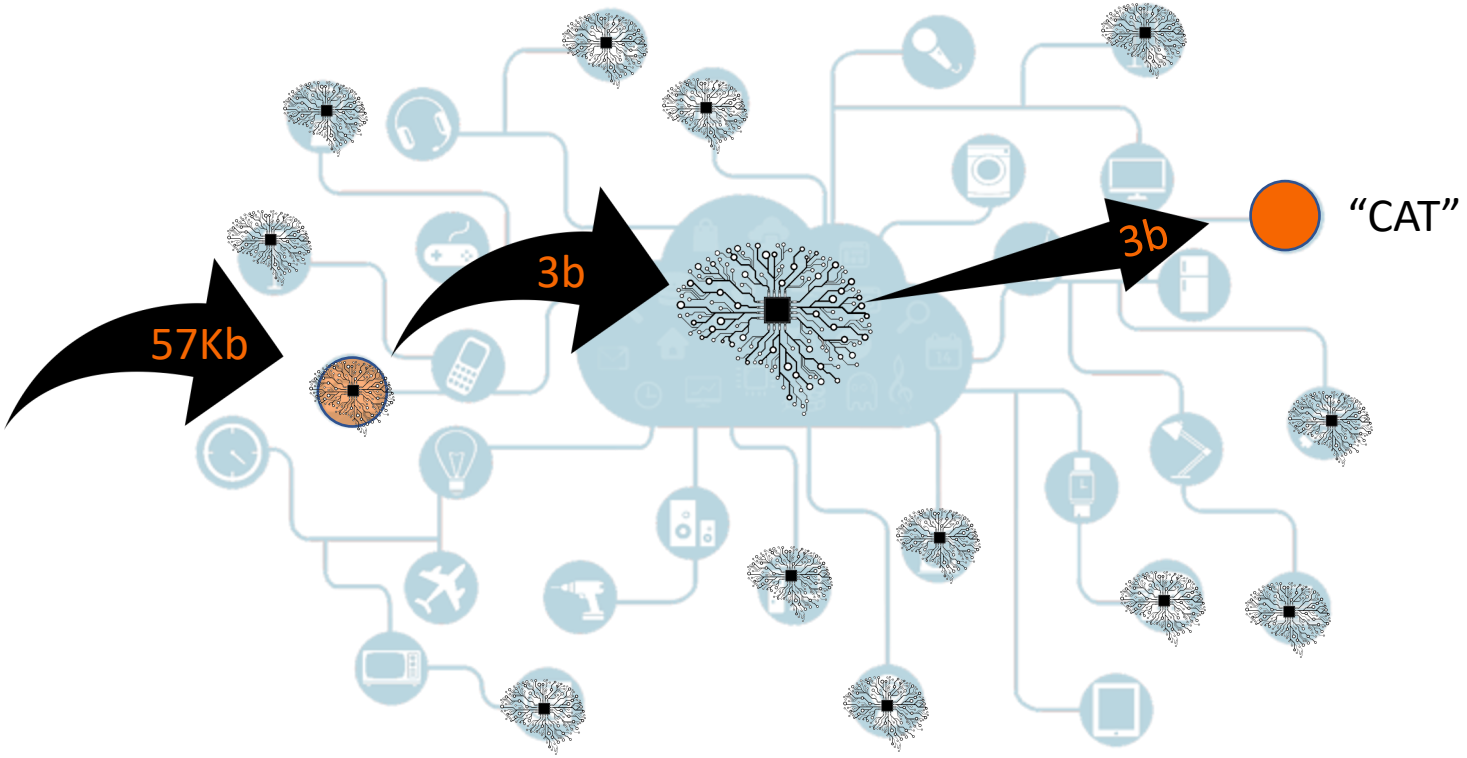


(146, 134, 3) ➔ 146 ∗ 134 ∗ 3 = 57,486 pixels

1 pixel ➔ 0 to 255 (1 byte)

AI at the Edge ➔ Edge AI

Meetup
IoT hacking & Data Science Innovators



"The term "Edge AI" might be the new buzzword of 2019/2020, much like "Internet of Things" was in 2016/2017. "

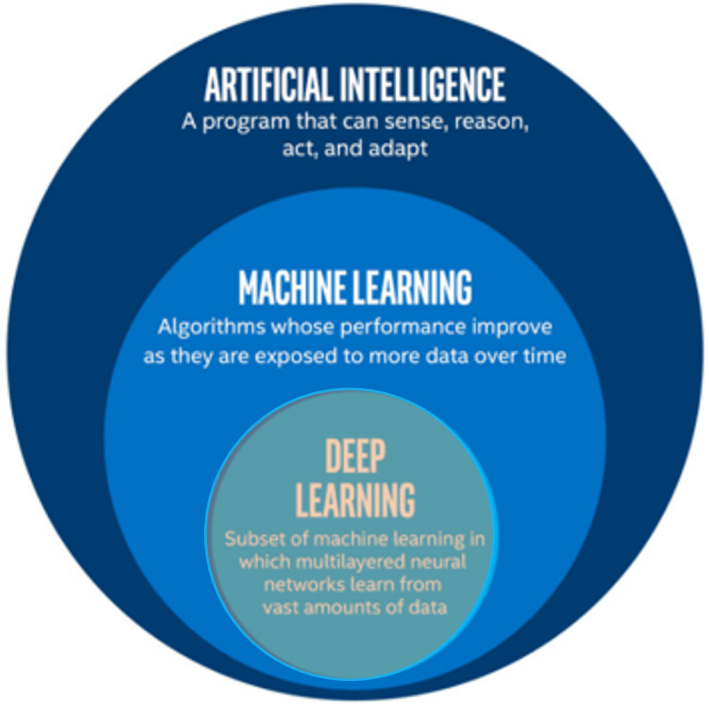
https://www.digikey.com/en/maker/projects/what-is-edge-ai-machine-learning-iot/4f655838138941138aaad62c170827af

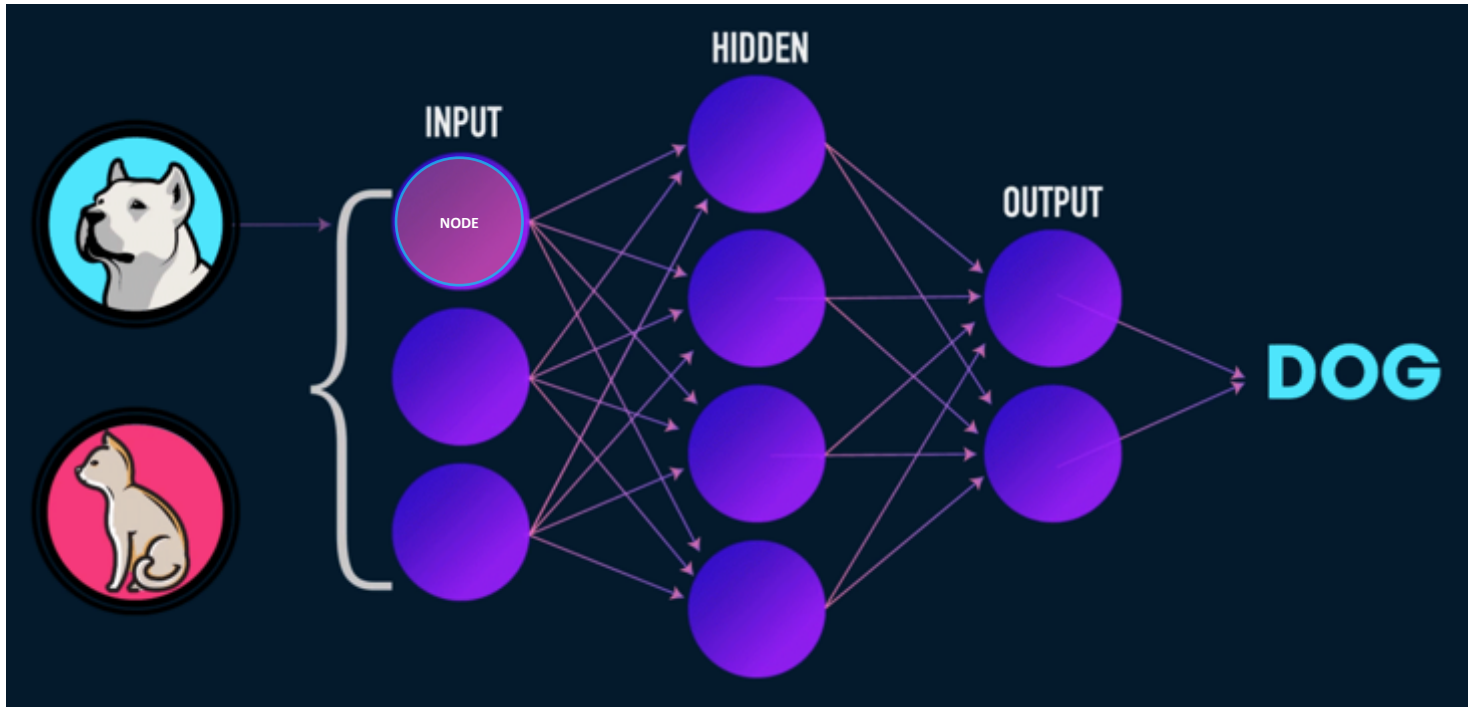
AI → Deep Learning (DL)

Meetup

IoT hacking & Data Science Innovators

Deep Learning: Subset of Machine Learning in which multilayered neural networks learn from vast amounts of data



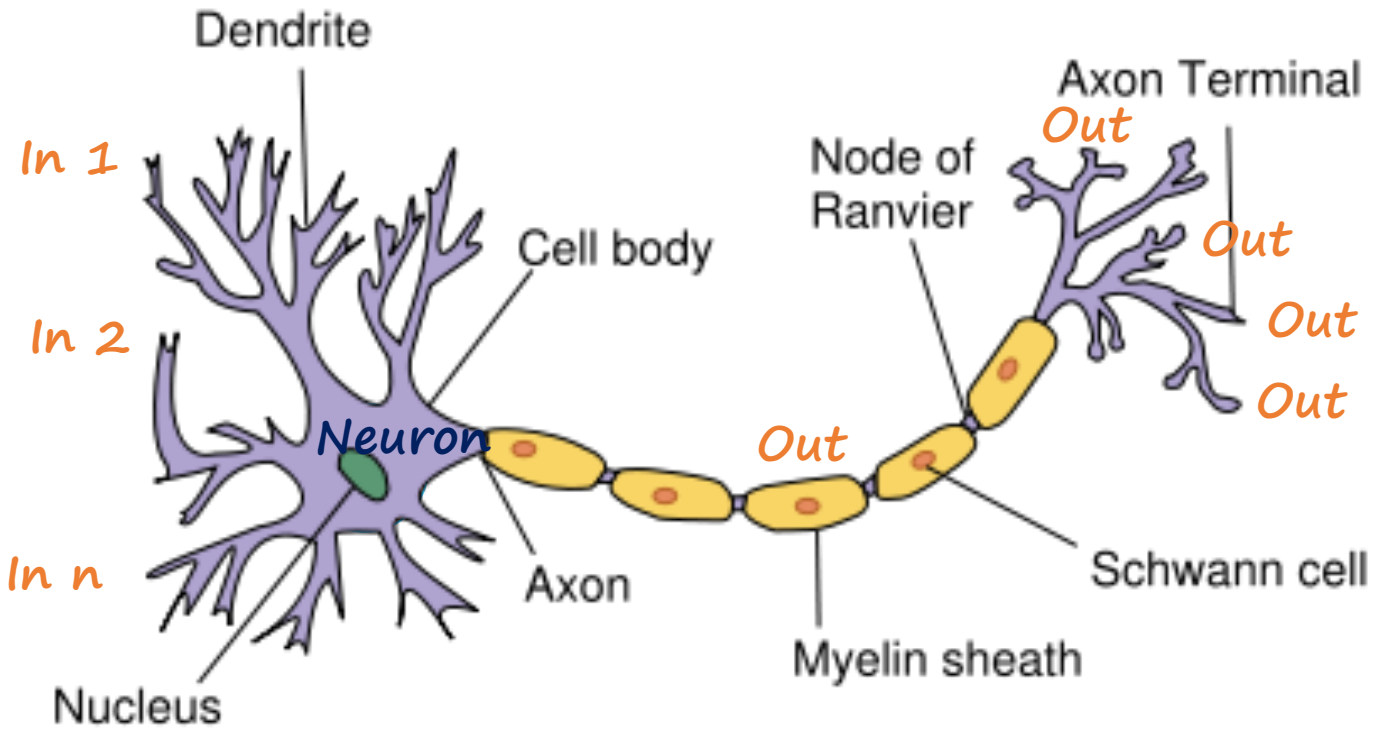


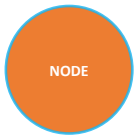
Source Image: https://towardsdatascience.com/an-introduction-to-deep-learning-af63448c122c

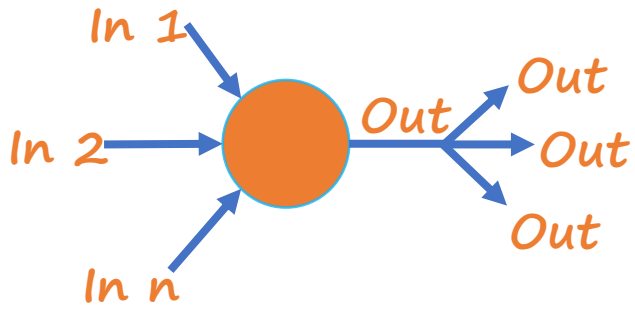
Neuron

Meetup

IoT hacking & Data Science Innovators

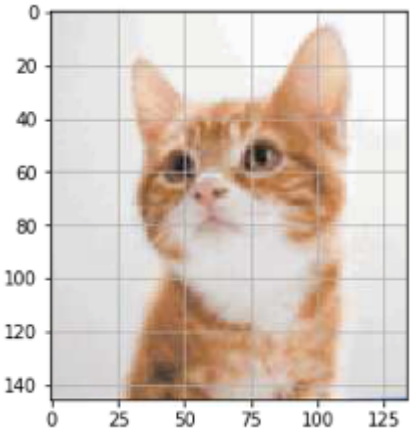






Artificial Neural Network

Meetup
IoT hacking & Data



(146, 134, 3)

RGB

```
[[227, 228, 227, ..., 249, 249, 249],
 [219, 219, 219, ..., 249, 249, 249],
 [221, 221, 221, ..., 249, 249, 249],
 ...,
 [209, 209, 209, ..., 220, 220, 218],
 [209, 209, 209, ..., 205, 205, 204],
 [208, 209, 209, ..., 124, 119, 121]]
```

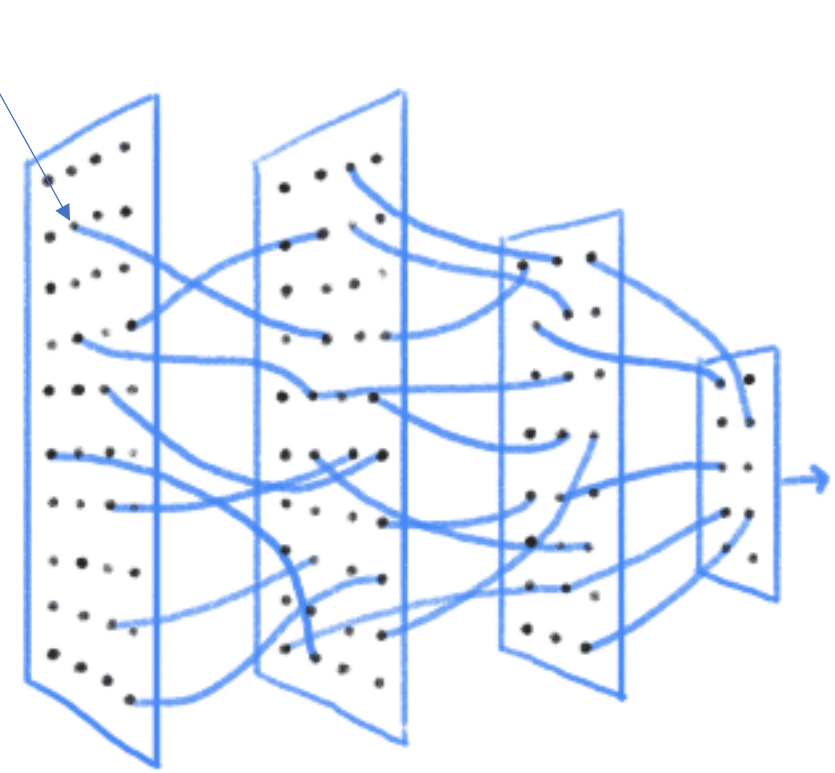
146 * 134 * 3 = 57,486 pixels

[227, 228, 227... 124, 119, 121]

CAT

(LABELED PHOTOS)

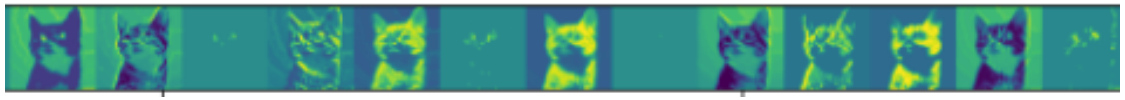
DOG



OUTPUT

cat

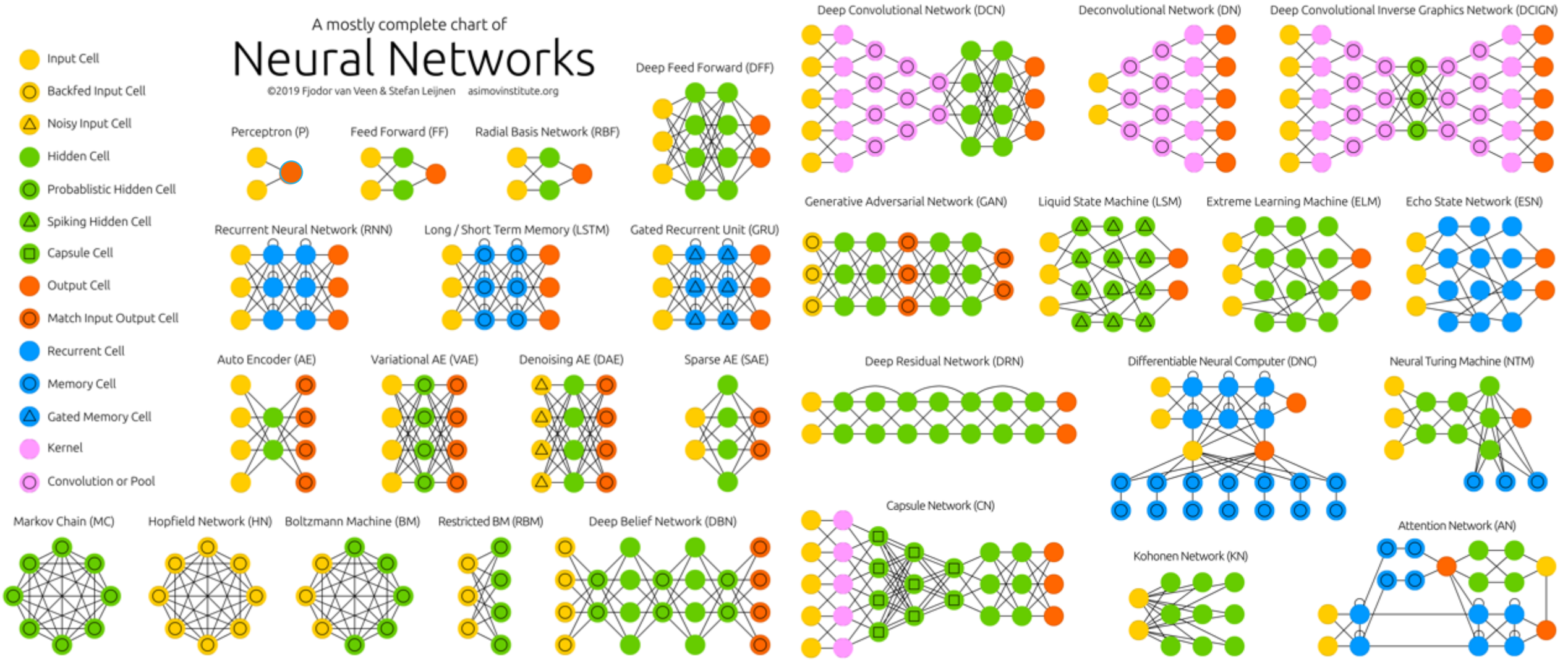
CNN (Layer 1) ➔



The Neural Network Zoo

Meetup

IoT hacking & Data Science Innovators

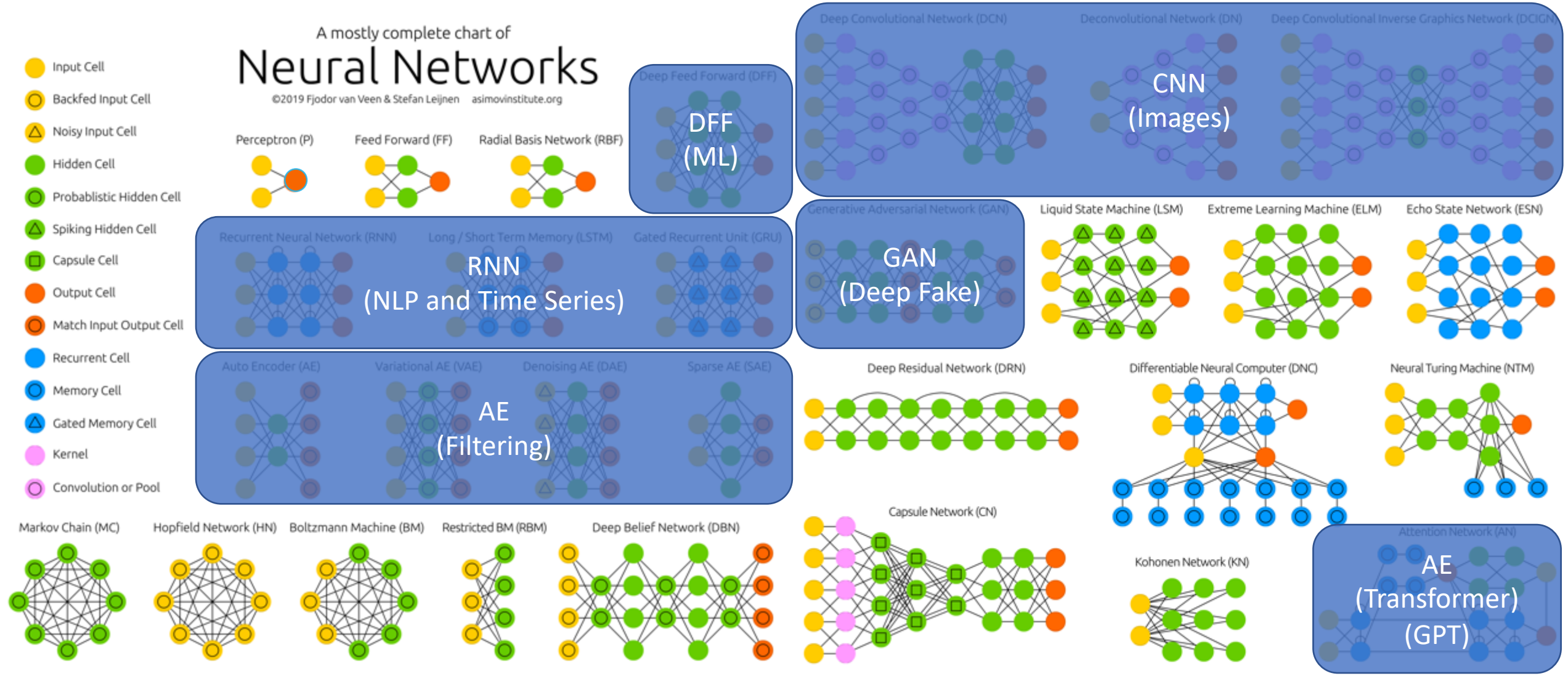


https://www.asimovinstitute.org/neural-network-zoo/

The Neural Network Zoo

Meetup

IoT hacking & Data Science Innovators

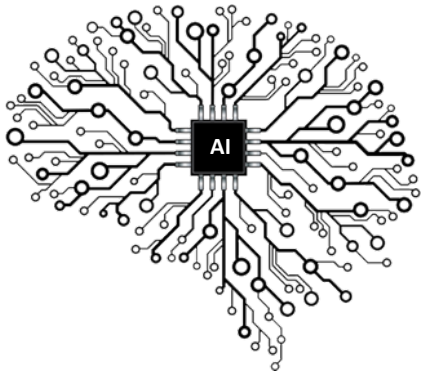


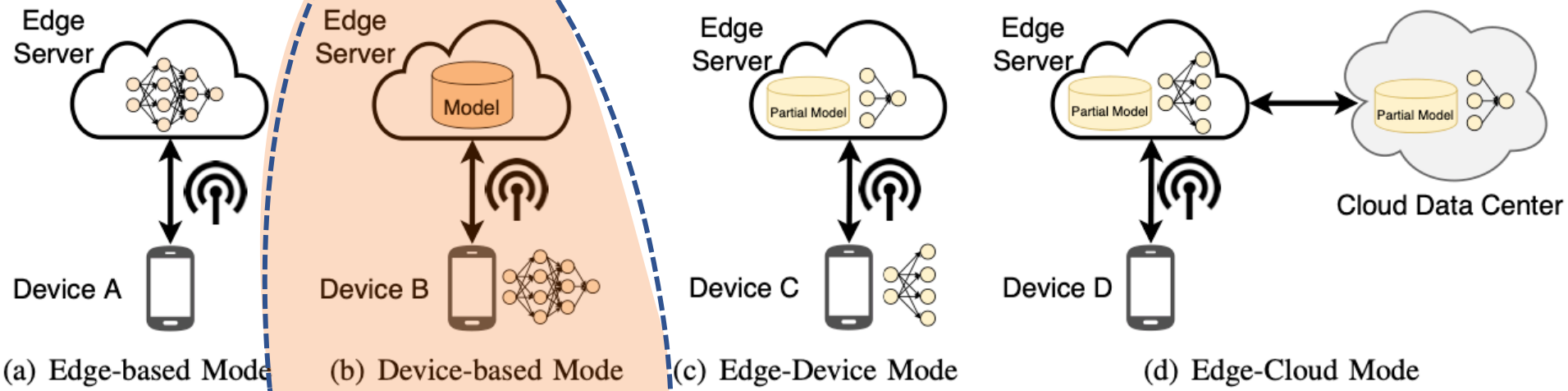
https://www.asimovinstitute.org/neural-network-zoo/

AI → Edge Computing

Meetup

IoT hacking & Data Science Innovators

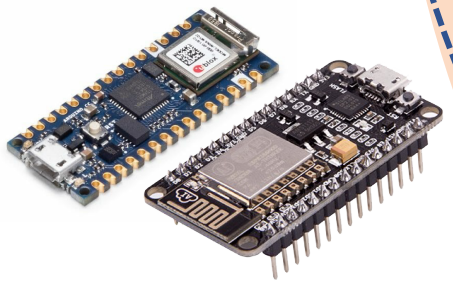




(a) Edge-based Mode (b) Device-based Mode (c) Edge-Device Mode (d) Edge-Cloud Mode

https://arxiv.org/pdf/1905.10083.pdf

Smart Phones



MCU (Nano 33IOT / ESP32)

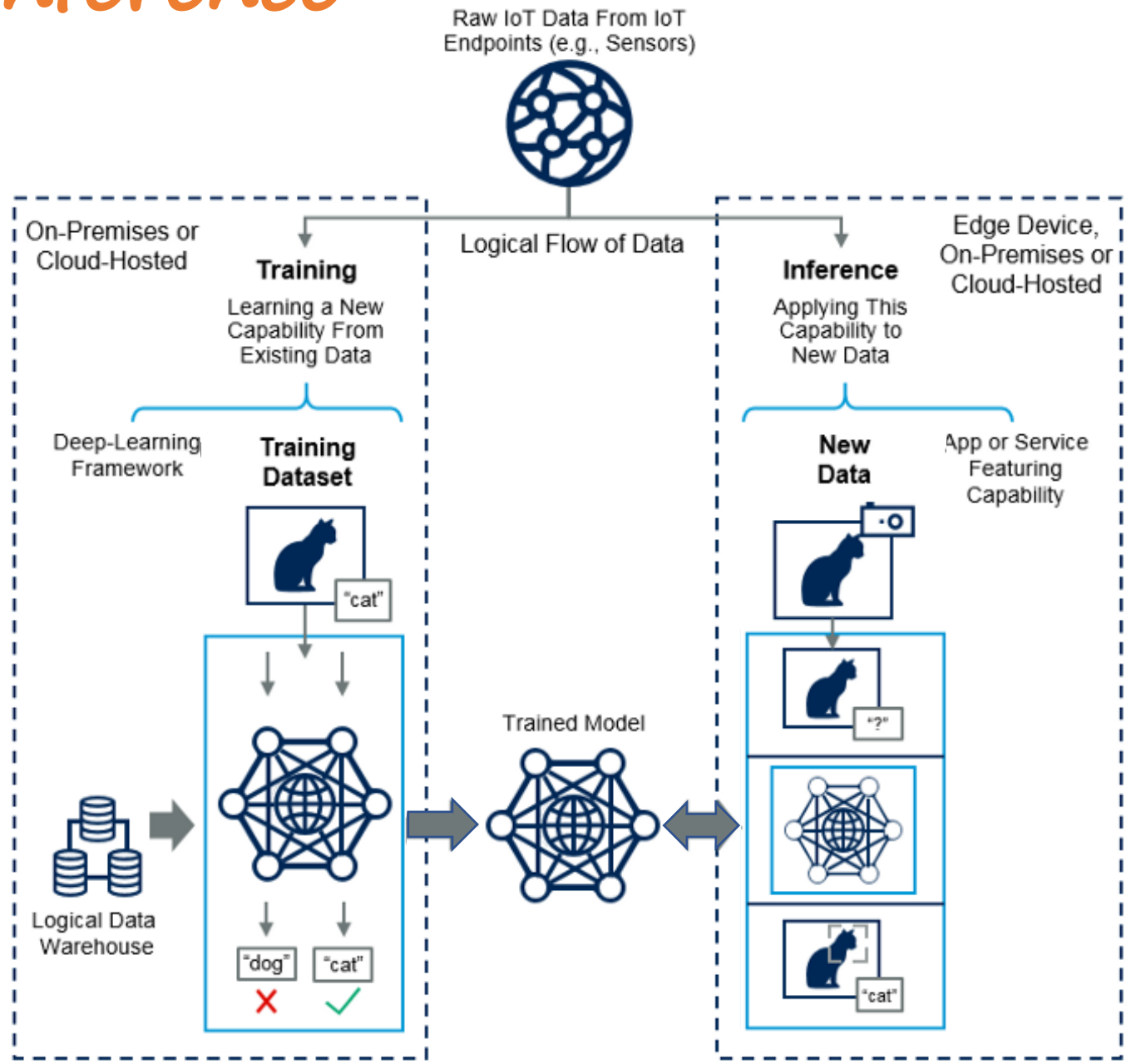
RaspBerry Pi

BeagleBone

NVIDIA Jetson

Training vs. Inference

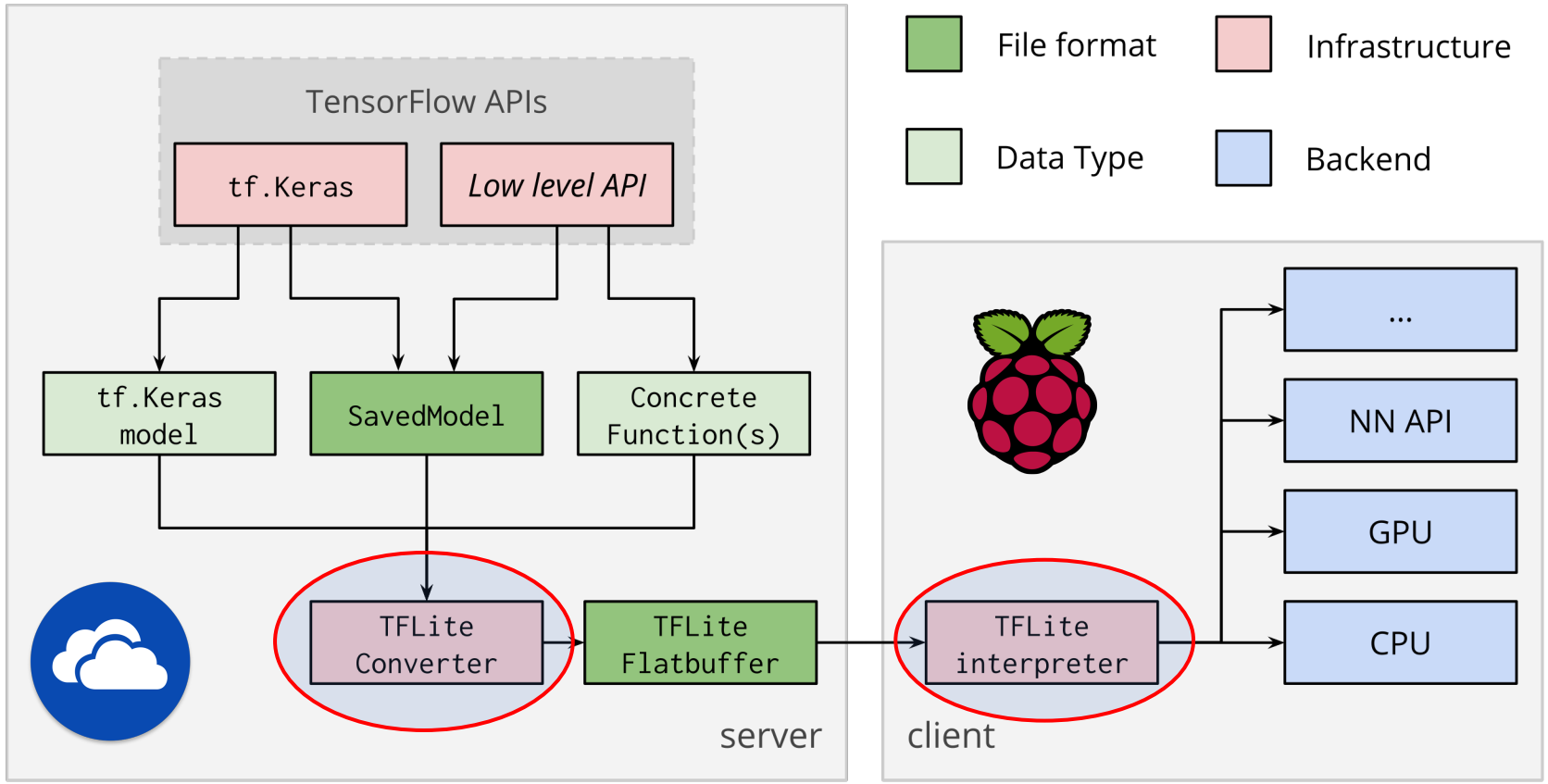
IoT hacking & Data Science Innovators



https://blogs.gartner.com/paul-debeasi/2019/02/14/training-versus-inference/

TensorFlow Lite

Meetup
IoT hacking & Data Science Innovators



Meetup

IoT hacking & Data Science Innovators

Image Classification & Object Detection Live DEMO

Deep Learning Workflow

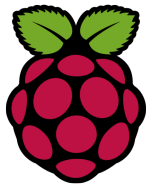
Meetup
IoT hacking & Data Science Innovators

Decide a Goal → Collect a dataset → Design a model architecture

On-Premises / Cloud (Server)

Train a model → Convert a model

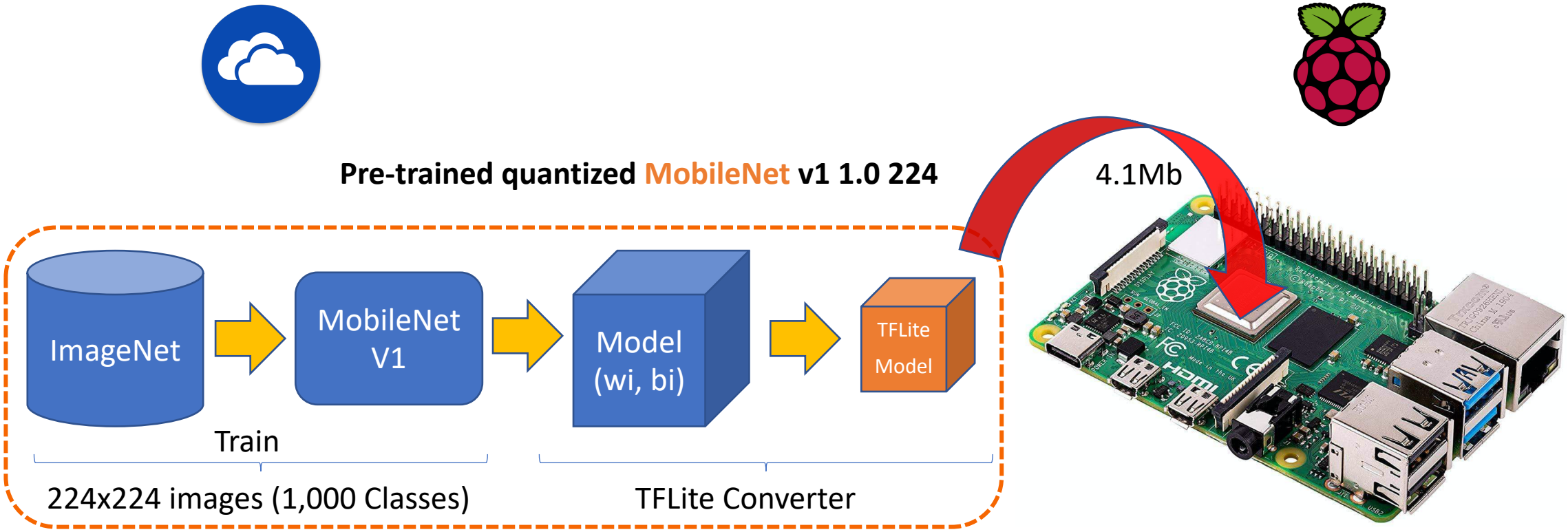
Edge Device (Client)



Run Inference → Evaluate and Troubleshoot

Demo 1 - Image Classification

Meetup
IoT hacking & Data Science Innovators



TF Lite Inference – Image Classification

Meetup
IoT hacking & Data Science Innovators

```
In [5]:  1  input_details

Out[5]: [{'name': 'input',
          'index': 88,
          'shape': array([  1, 224, 224,   3]),
          'dtype': numpy.uint8,
          'quantization': (0.0078125, 128),
          'quantization_parameters': {'scales': array([0.0078125], dtype=float32),
           'zero_points': array([128]),
           'quantized_dimension': 0}}]
```

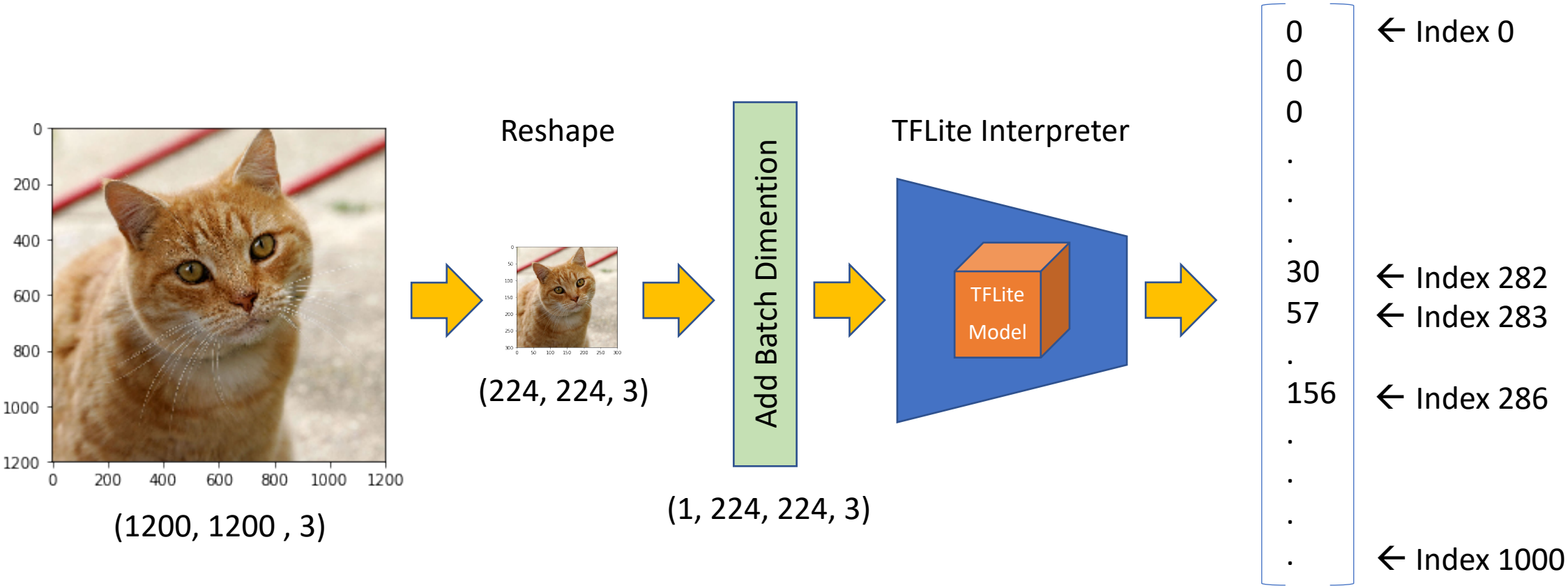
Input Image Shape

```
In [6]:  1  output_details

Out[6]: [{'name': 'MobilenetV1/Predictions/Reshape_1',
          'index': 87,
          'shape': array([   1, 1001]),
          'dtype': numpy.uint8,
          'quantization': (0.00390625, 0),
          'quantization_parameters': {'scales': array([0.00390625], dtype=float32),
           'zero_points': array([0]),
           'quantized_dimension': 0}}]
```

Output model

TF Lite Inference – Image Classification

Meetup
IoT hacking & Data Science Innovators

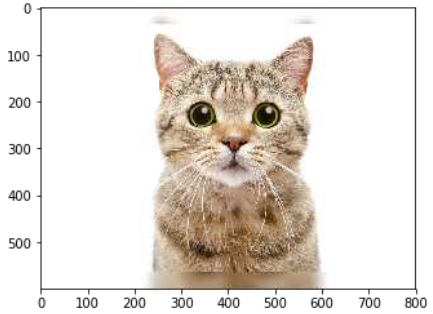


TF Lite Inference – Demo - Results

IoT hacking & Data Science Innovators

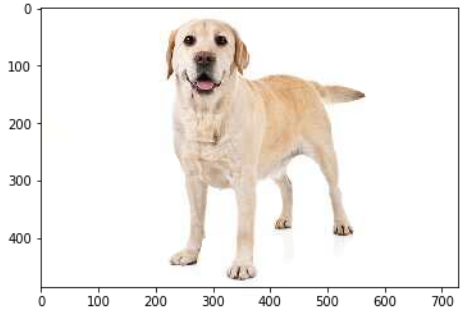
```
image_path = './images/cat_1.jpg'
image_classification(image_path, labels)
```

```
[PREDICTION]        [Prob]

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



```
image_path = './images/dog_1.jpg'
image_classification(image_path, labels)
```

```
[PREDICTION]        [Prob]

Labrador retriever  : 83%
golden retriever    : 13%
bloodhound          : 0%
```



```
image_path = './images/pi_1.jpeg'
image_classification(image_path, labels)
```

```
[PREDICTION]        [Prob]

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

```
image_path = './images/car_2.jpg'
image_classification(image_path, labels)
```

```
[PREDICTION]        [Prob]

minivan             : 76%
convertible         : 6%
jeep                : 4%
```

```
image_path = './images/ship_1.jpg'
image_classification(image_path, labels)
```

```
[PREDICTION]        [Prob]

liner               : 85%
trimaran            : 11%
aircraft carrier    : 0%
```

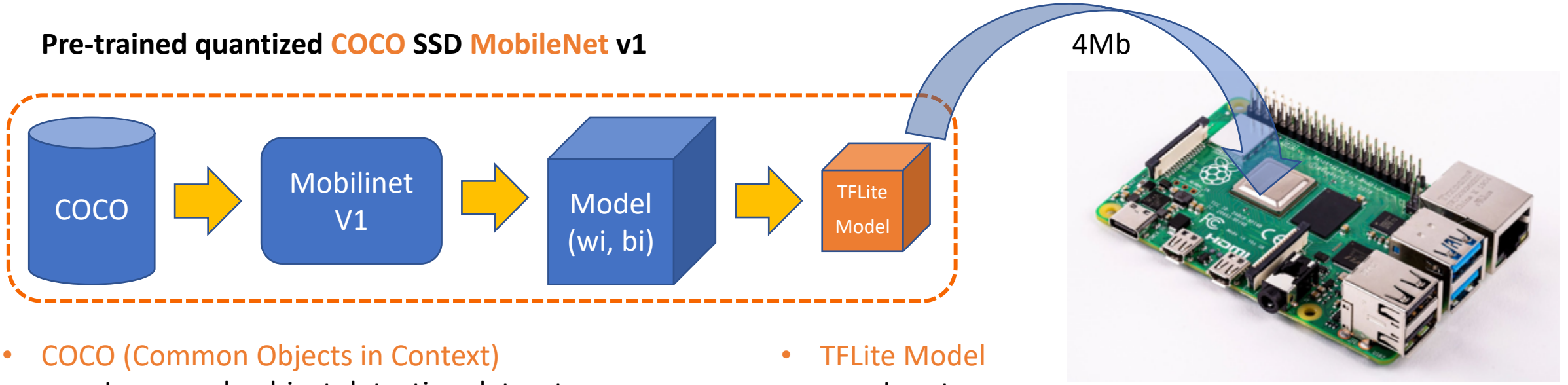
```
image_path = './images/table.jpg'
image_classification(image_path, labels)
```

```
[PREDICTION]        [Prob]

dining table        : 40%
table lamp          : 34%
desk                : 17%
```

Demo 2 - Object Detection

Meetup
IoT hacking & Data Science Innovators

Pre-trained quantized COCO SSD MobileNet v1



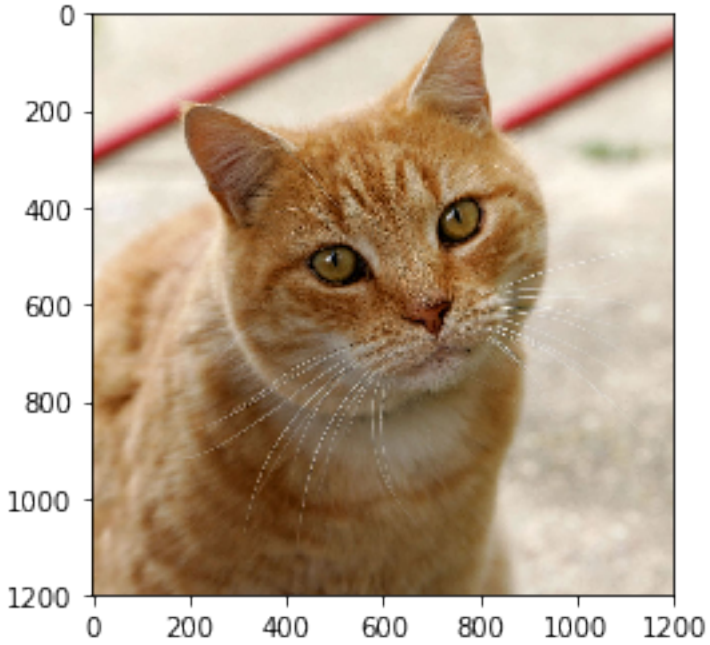
- COCO (Common Objects in Context)
  - Large-scale object detection dataset
  - 200K labeled images
  - 91 Stuff categories
- MobileNet
  - “Depth- wise Separable” convolutions
  - Introduced by Google in 2017
  - Similar performance with state-of-the-art architectures (as VGG or Inception)
  - Much smaller network (20% of VGG parameters (7M)
- TFLite Model
  - Input:
    - Image: 300x300x3
    - Flattened Buffer: 270K bytes
    - Each byte: 0 to 255
  - Output:
    - Bounding Box ➔ (ymin, xmin, ymax, xmax)
    - Class ID (0 to 89) ("Stuff categories")
    - Score (0 to 1)

COCO Paper: https://arxiv.org/pdf/1405.0312.pdf

MobileNet paper: https://arxiv.org/pdf/1704.04861.pdf

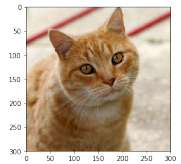
TF Lite Inference - Object Detection

Meetup
IoT hacking & Data Science Innovators



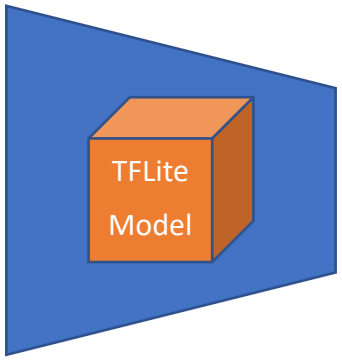
(1200, 1200, 3)

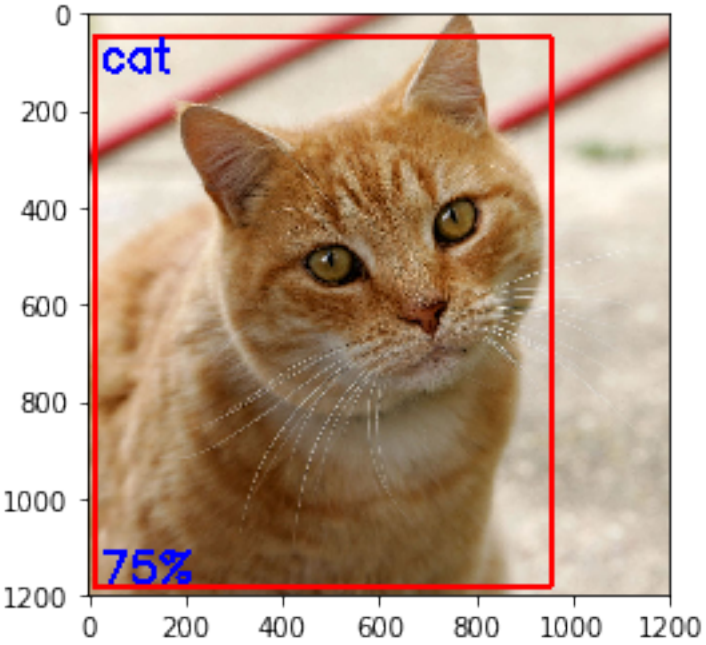
Pre-Process



(1, 300, 300, 3)

TFLite Interpreter

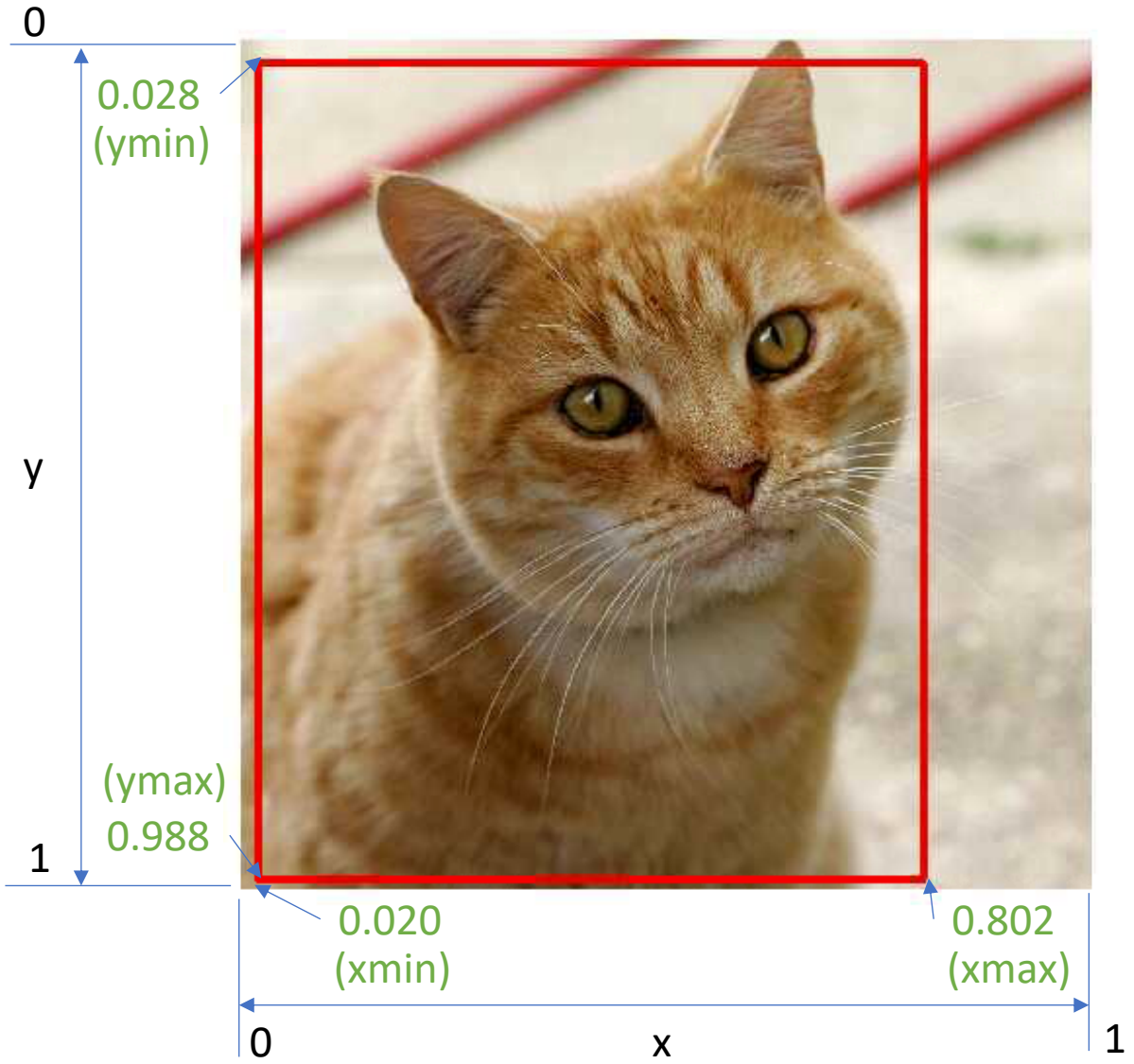




[{'bounding_box': [0.04125446,
0.01441014,
0.9867869 ,
0.79658735]
'class_id': 16.0,
'score': 0.7578125}]

Bounding Box- Object Detection

Meetup
IoT hacking & Data Science Innovators

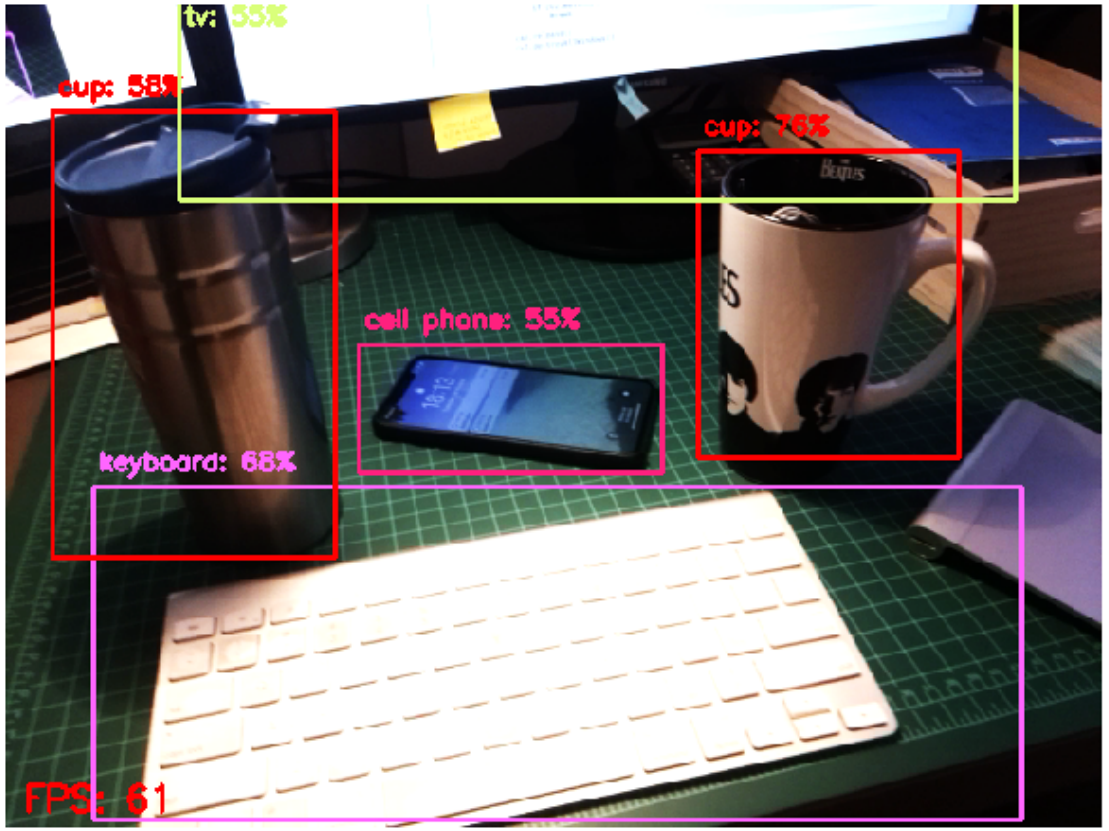


TF Lite Inference – Demo - Results

Meetup

IoT hacking & Data Science Innovators

Photos

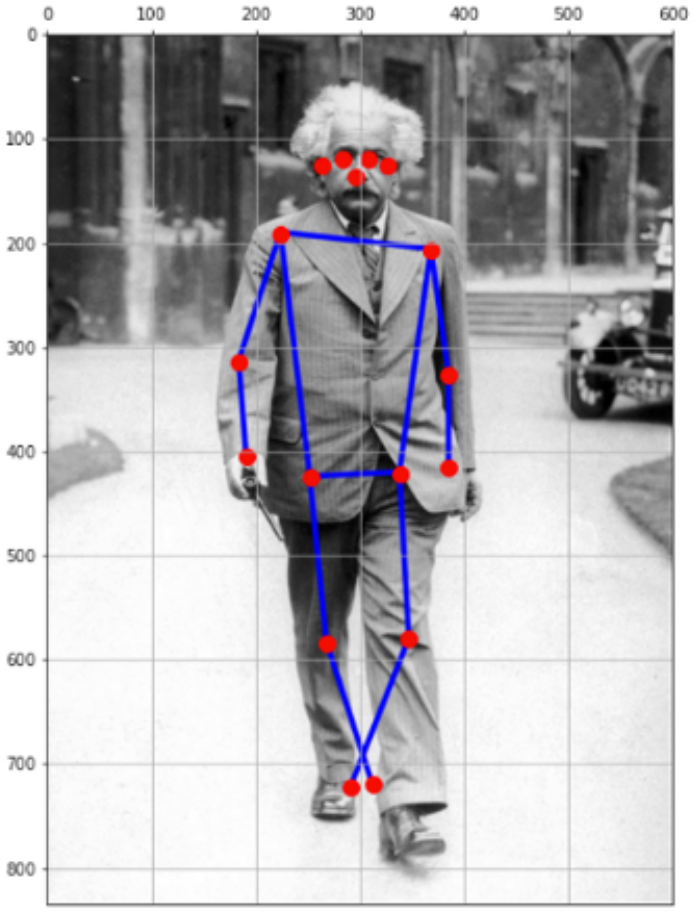


Live Video

TF Lite Inference – Other applications

Meetup
IoT hacking & Data Science Innovators

Pose Estimation on RPi



TinyML

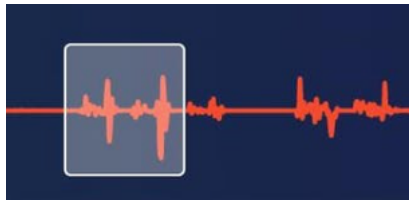
– Machine Learning on Ultra-Low-Power Microcontrollers

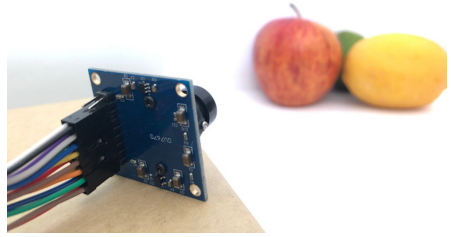
The world of Embedded devices:

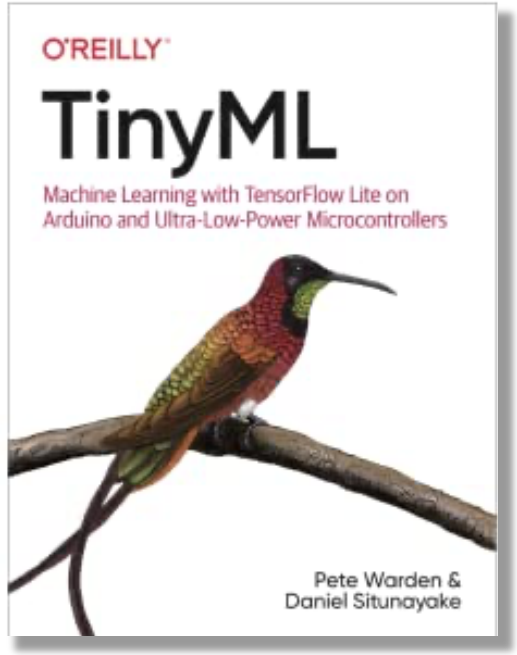
- Run a neural network model at an energy cost of below 1 mW
- Devices that can run on coin battery by a year
- 32 bit Microcontrollers that cost less than $1
- ARM's Cortex-M series of chips

Apllications as:

- Speech Recognition (Wake-Word Detection)
- Detecting Gestures (iWatch)
- Detection of Objects on camera







Learning more about Edge AI

Meetup
IoT hacking & Data Science Innovators

- **Deploy machine learning models on mobile and IoT devices:**
  - https://www.tensorflow.org/lite

- **Teachable Machine**: Train a computer to recognize your own images, sounds, & poses:
  - https://teachablemachine.withgoogle.com

- **Machine Learning on Mobile and Edge Devices with TensorFlow Lite: Daniel Situnayake at QCon SF**
  - https://www.infoq.com/news/2019/11/tensorflow-lite-edge-qconsf/

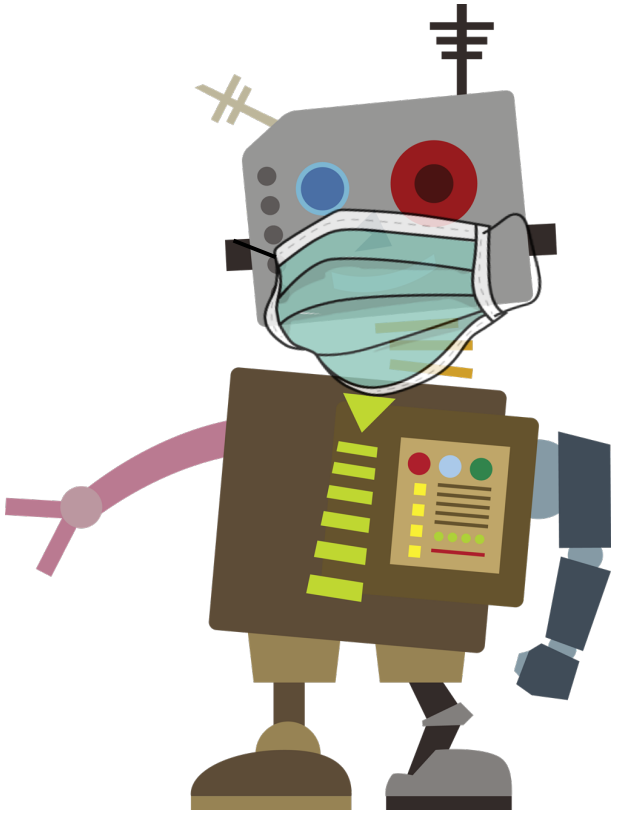
- **What is Edge AI? Machine Learning + IoT**
  - https://www.digikey.com/en/maker/projects/what-is-edge-ai-machine-learning-iot/4f6558313894113 8aaad62c170827af

- **Exploring AI at the Edge!**
  - https://towardsdatascience.com/exploring-ia-at-the-edge-b30a550456db

Meetup

IoT hacking & Data Science Innovators

"To infinity and beyond!"



MJRoBot.org

github.com/Mjrovai

hackster.io/mjrobot

medium.com/@rovai

instructables.com/member/mjrovai

Meetup

IoT hacking & Data Science Innovators

Thanks

And keep safe!