

IOT_2: Unleash the power of IoT data

protocols, analysis, artificial intelligence, machine learning,...



Capsule Booster – 2022

Prof. Congduc Pham
<http://www.univ-pau.fr/~cpham>



Horizon 2020
European Union funding
for Research & Innovation



Advanced and disruptive IoT/AI technologies targeting the smallholder community for increased resilience



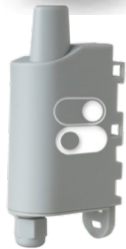
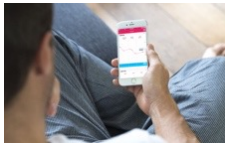
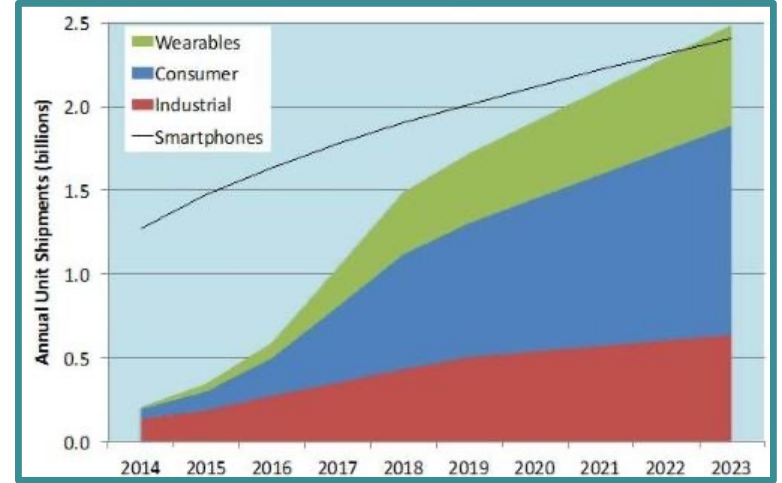


IOT

TECHNOLOGY ?
CONCEPT ?

IoT=communicating objects

Pr. Congduc Pham
/~cpham
Poids (kg)
TARIFOND 79.6kg
NOUVEAU TRAV. 72.2kg
<http://>



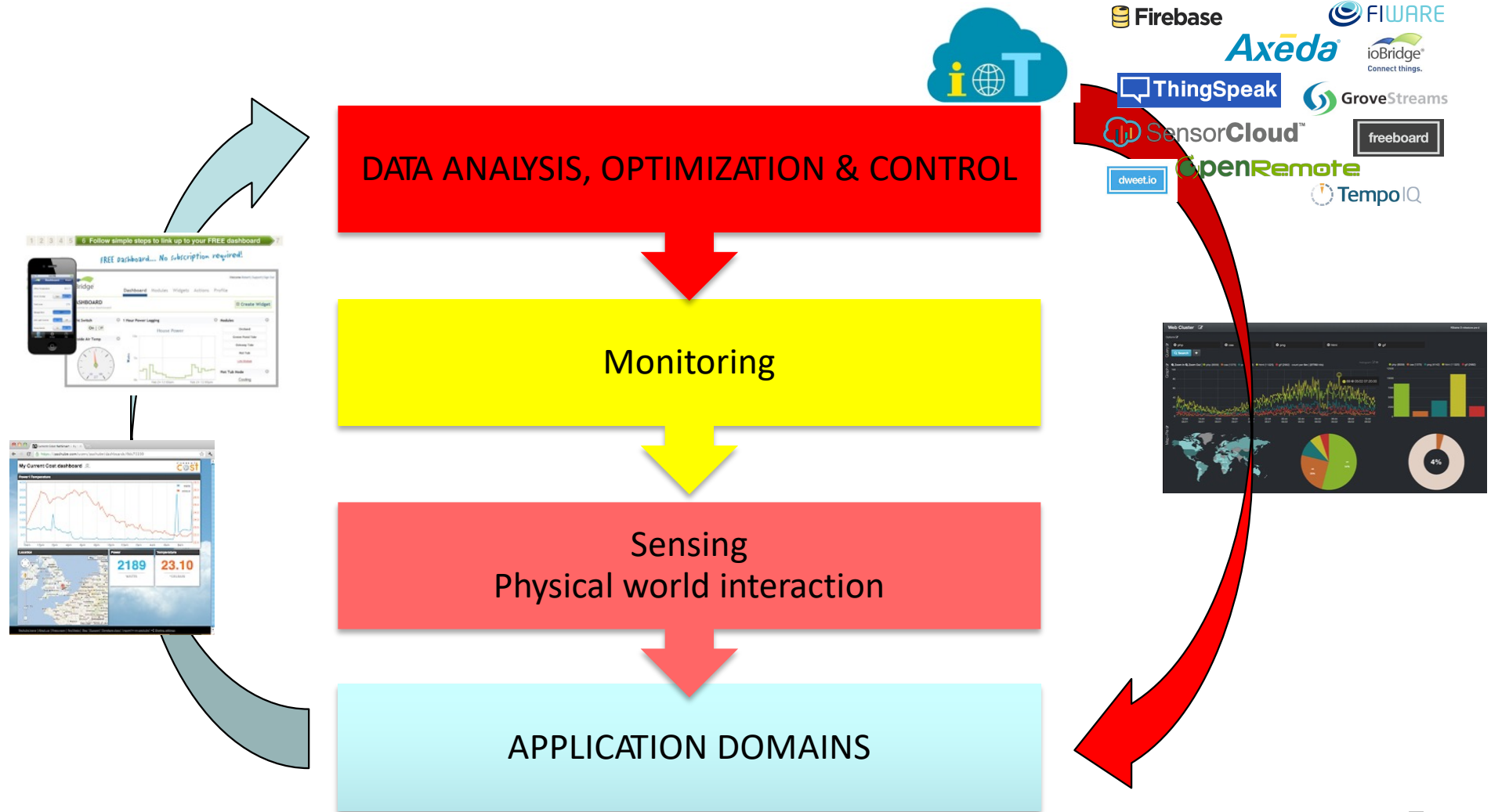
IoT=interactions with physical world



Pr. Congduc Pham
http://www.univ-pau.fr/~cpham

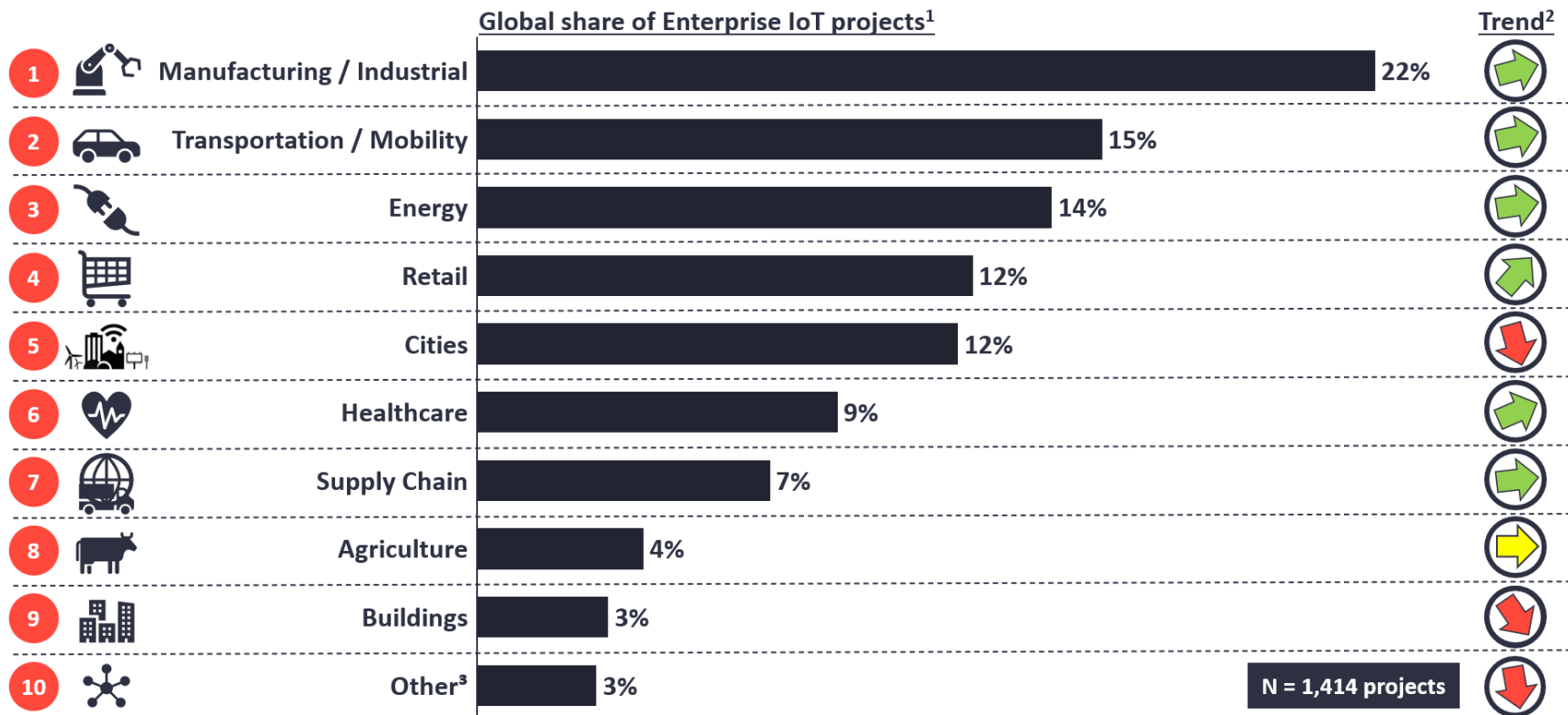


Sense, Monitor, Optimize & Control



Top IoT applications, 2020

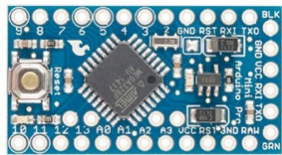
Top 10 IoT Application areas 2020



Note: 1. Based on 1,414 publically known IoT projects (not including consumer IoT projects eg smart home, wearables, etc.) 2. Trend based on relative comparison with % of projects in the 2018 IoT Analytics IoT project list e.g., a downward arrow means the relative share of all projects has declined, not the overall number of projects. 3. Other includes IoT projects from Enterprise & Finance sectors. **Source:** IoT Analytics Research - July 2020

Low-cost microcontroller boards

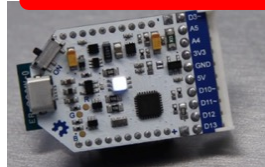
Arduino Pro Mini



LoPy

<http://blog.atmel.com/2015/12/16/rewind-50-of-the-best-boards-from-2015/>

<http://blog.atmel.com/2015/04/09/25-dev-boards-to-help-you-get-started-on-your-next-iot-project/>



Theairboard



Expressif ESP32



Teensy 3.2



LinkIt Smart7688 duo

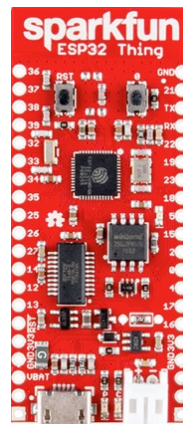
STM32 Nucleo-32



Heltec ESP32 + OLED



Adafruit Feather



Sparkfun ESP32 Thing



Tessel

SodaqOnev2

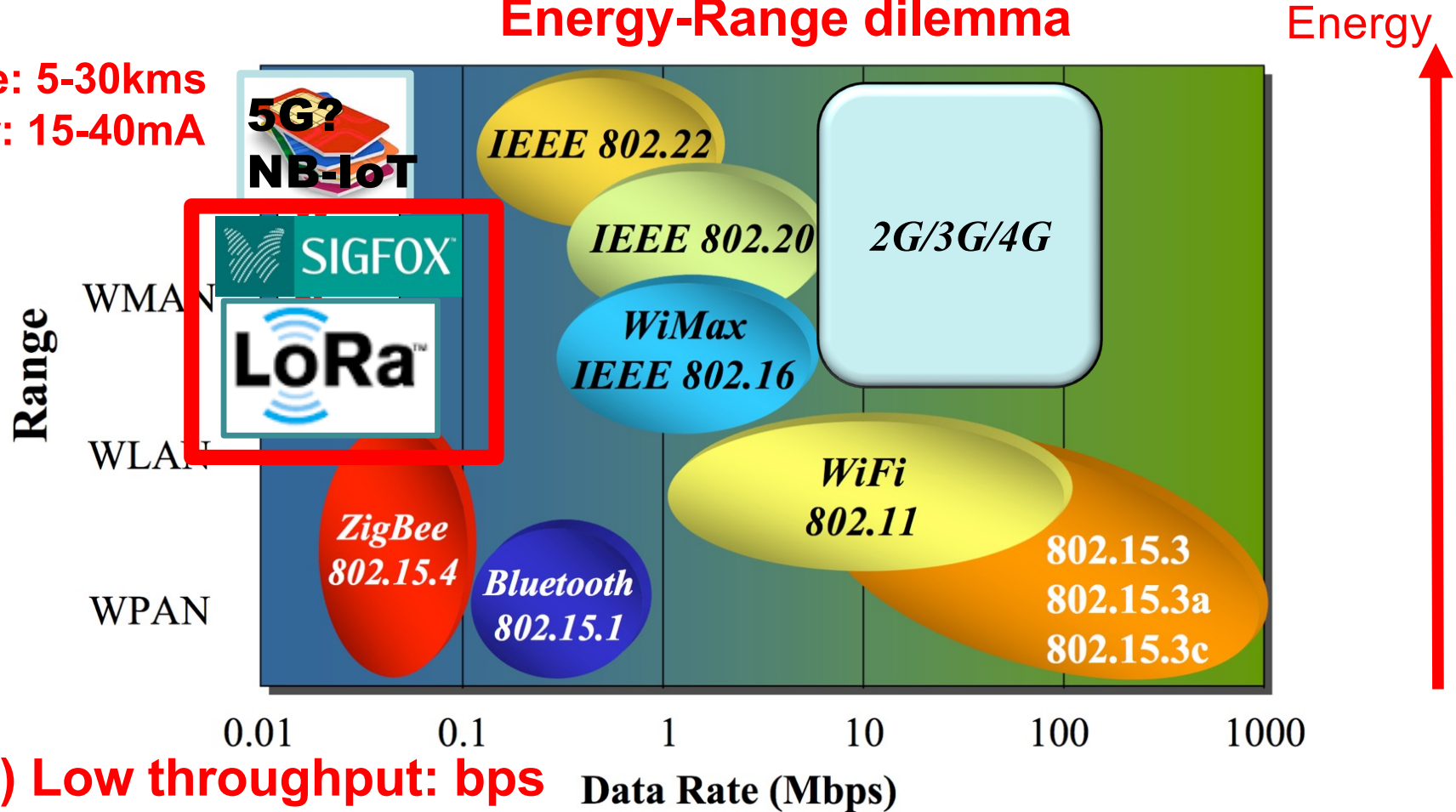


Tinyduino

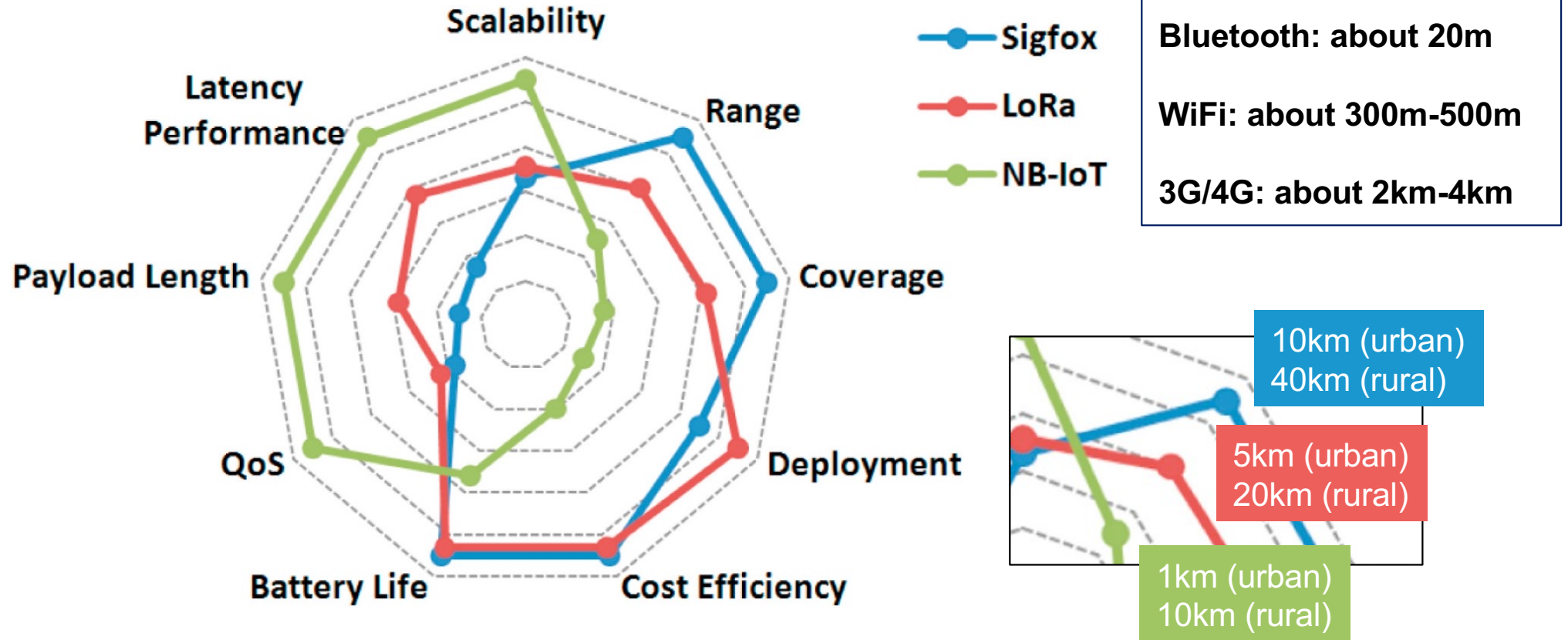
Low-power & long-range radios

Energy-Range dilemma

Long-range: 5-30kms
Low-power: 15-40mA



LPWAN expected range?



Pr. Congduc Pham
http://www.univ-pau.fr/~cpham

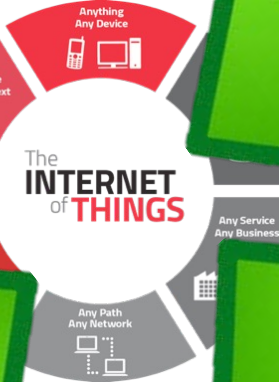
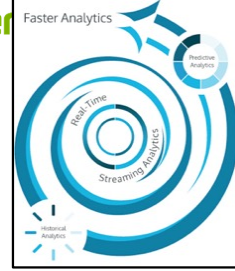
IoT becomes reality!

Boster 17 years later: the incredibly large microcontroller board ecosystem!

Pr. Congduc Pham
http://www.univ-pau.fr/~cpham

Boster LoRa modules with Semtech's SX

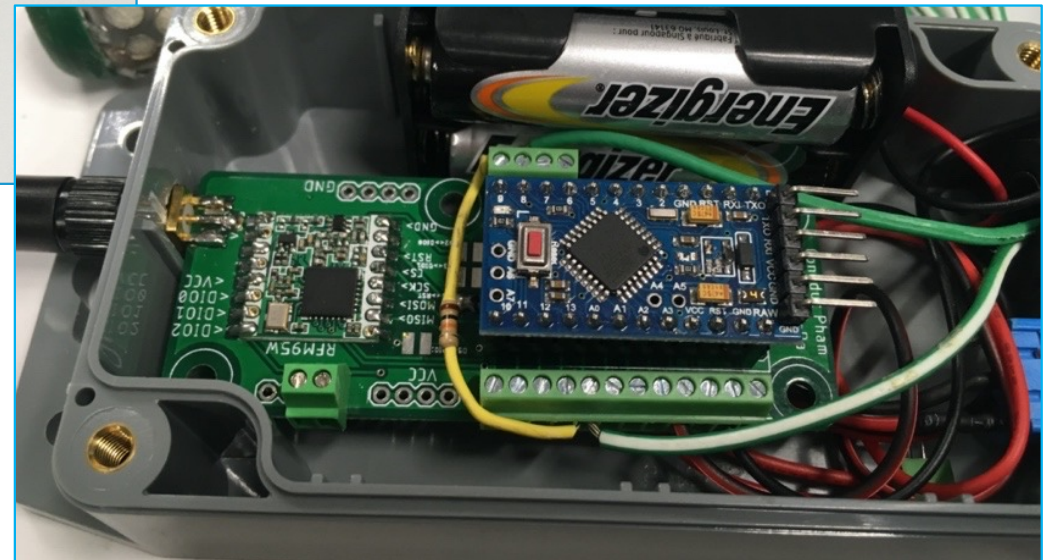
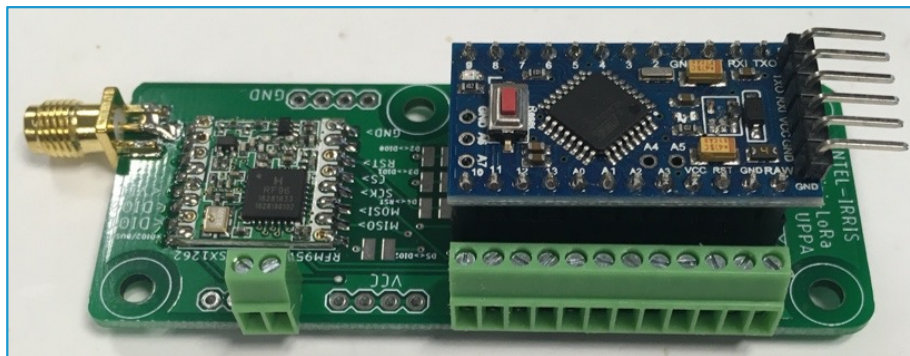
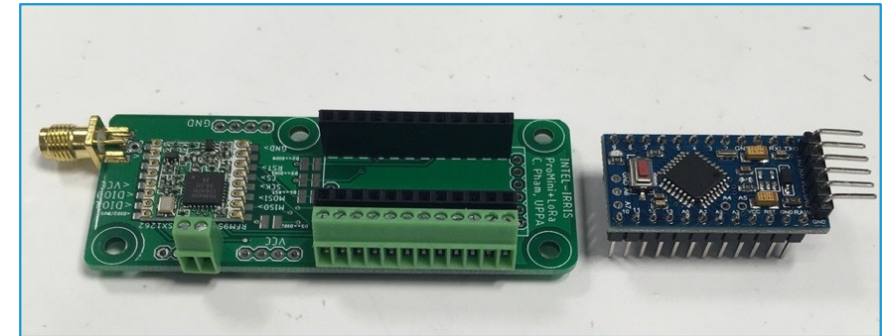
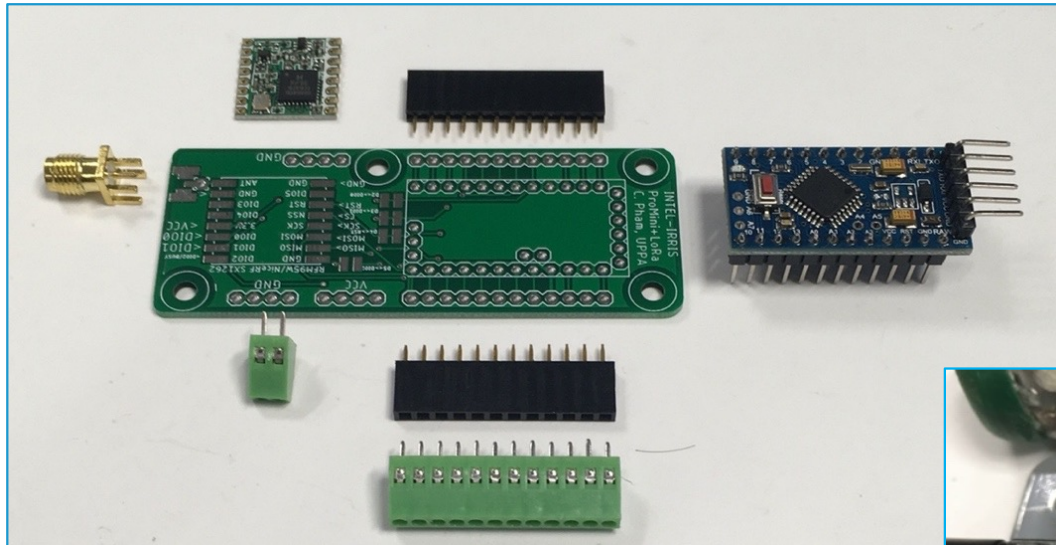
Pr. Congduc Pham
http://www.univ-pau.fr/~cpham

Pr. Congduc Pham
http://www.univ-pau.fr/~cpham

Simple & Modular design

- Simple integration on PCB of off-the-shelves components

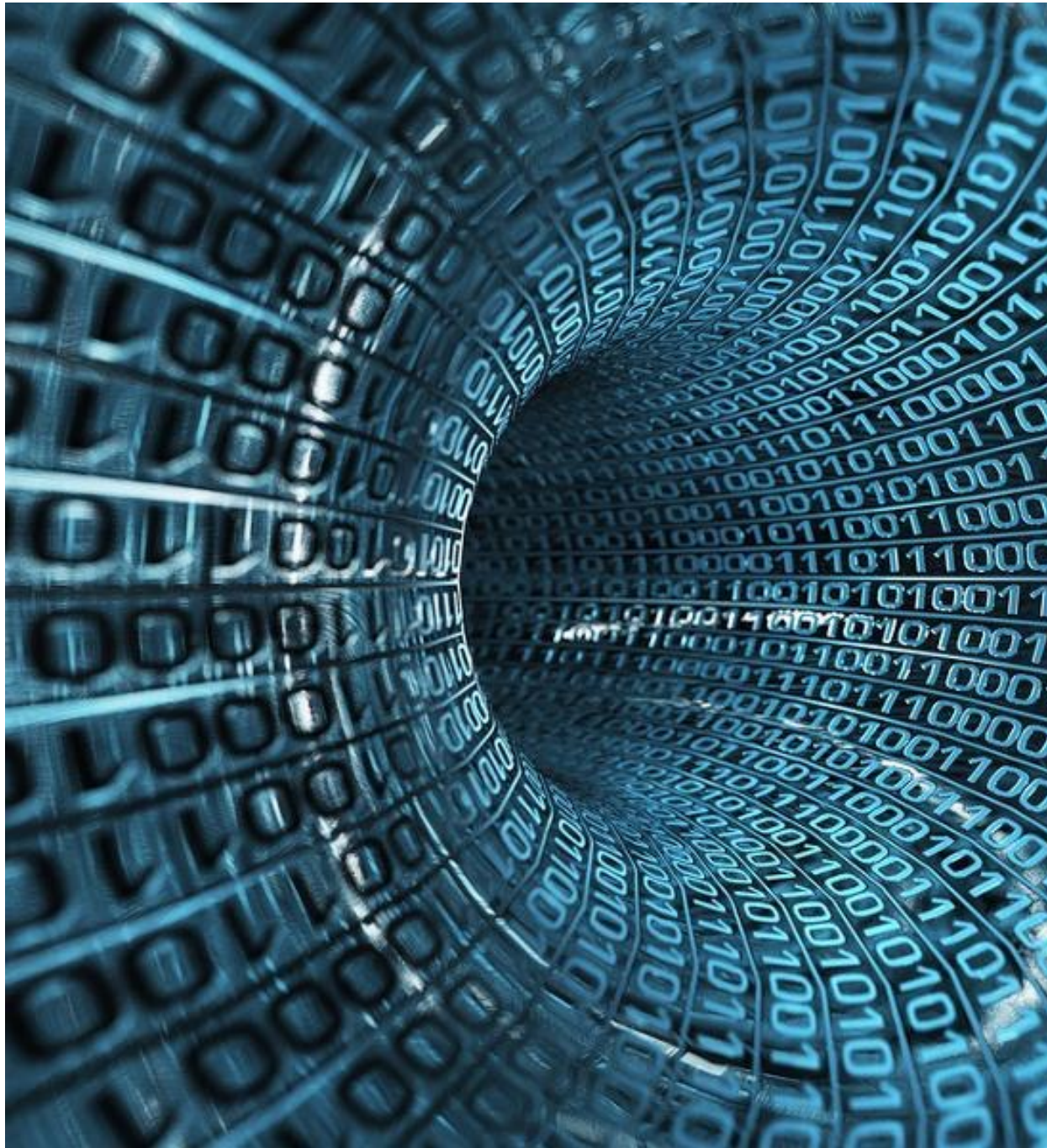


INTEL-IRRIS starter-kit

- ⦿ "Intelligent Irrigation in-the-box", "plug-&-sense", fully autonomous
- ⦿ From idea to reality!

Pr. Congduc Pham
<http://www.univ-pau.fr/~cpham>





IOT

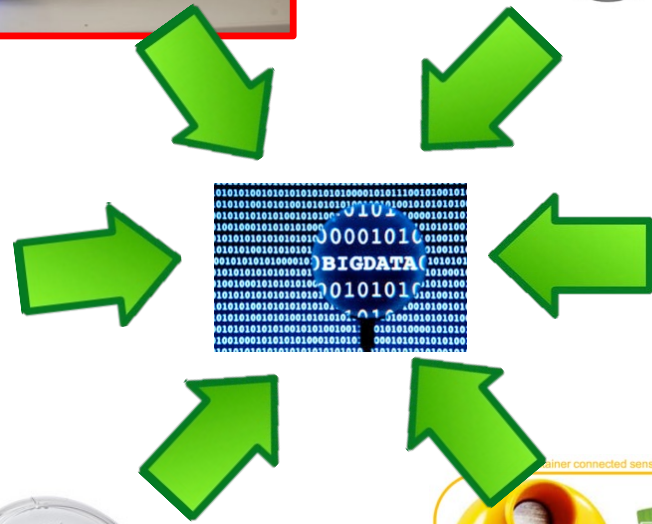
Unleash the power of
IoT data!



**2022, billions of IoT devices
are deployed worldwide!**

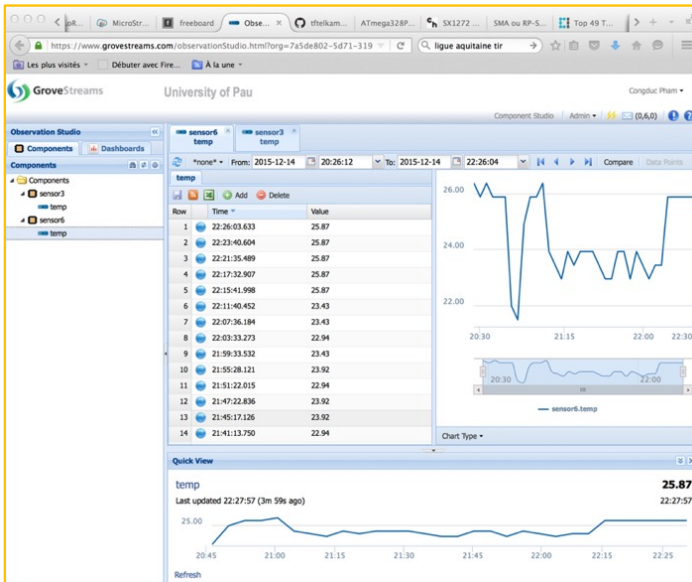
These things talk a lot!

Lot's of data !



Things talks to IoT clouds

Pr. Congduc Pham
<http://www.univ-pau.fr/~cpham>

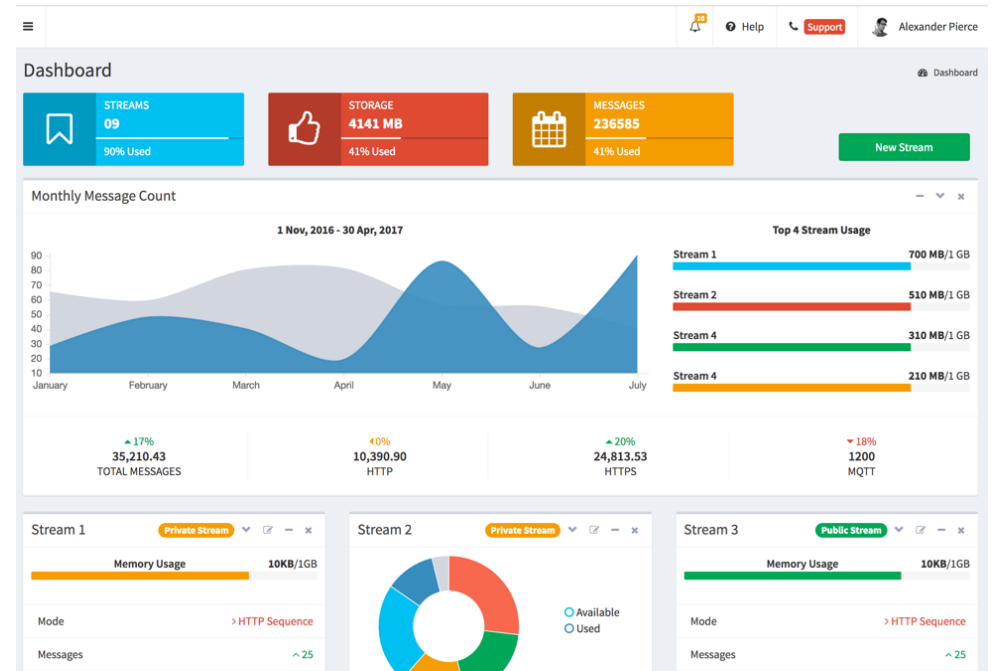
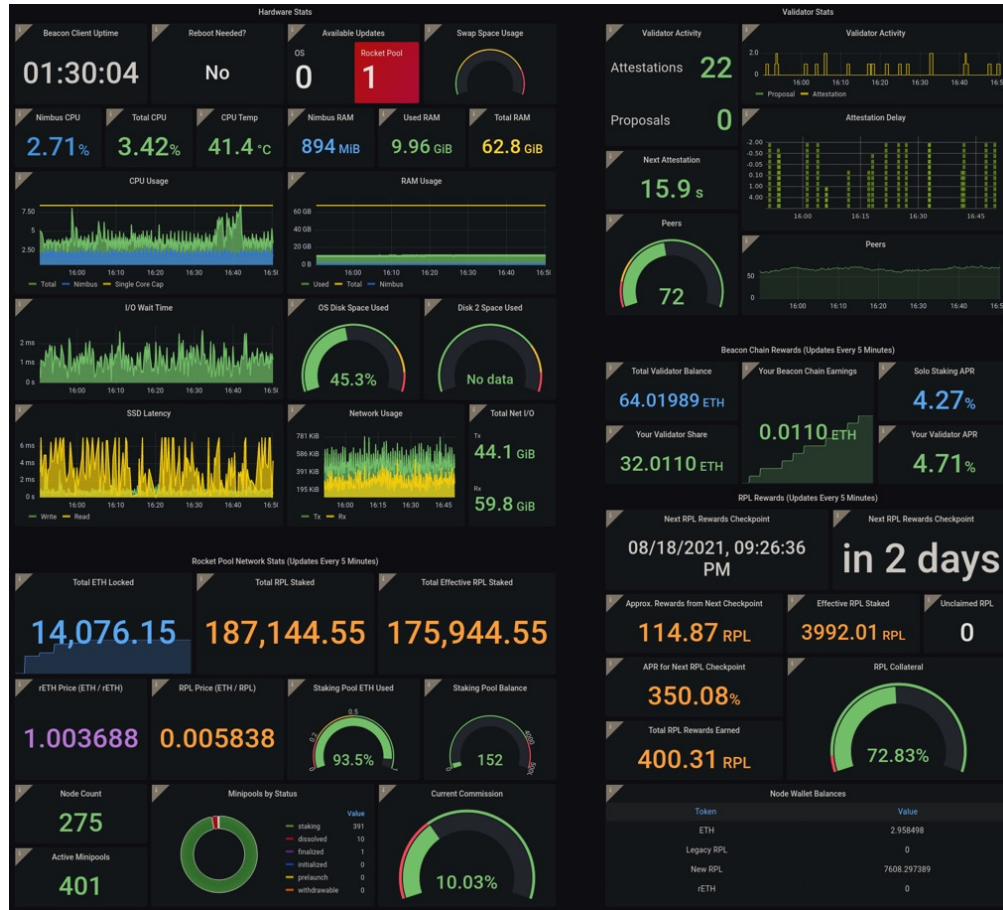


IoT added-values come from interactions and linked data!



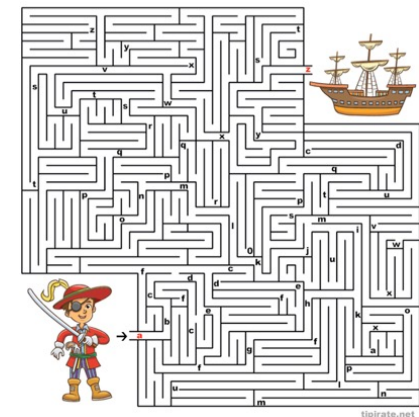
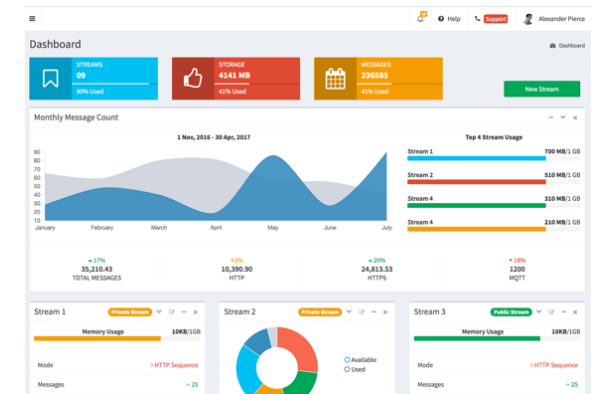
Integrating multiple data sources

Pr. Congduc Pham
http://www.univ-pau.fr/~cpham



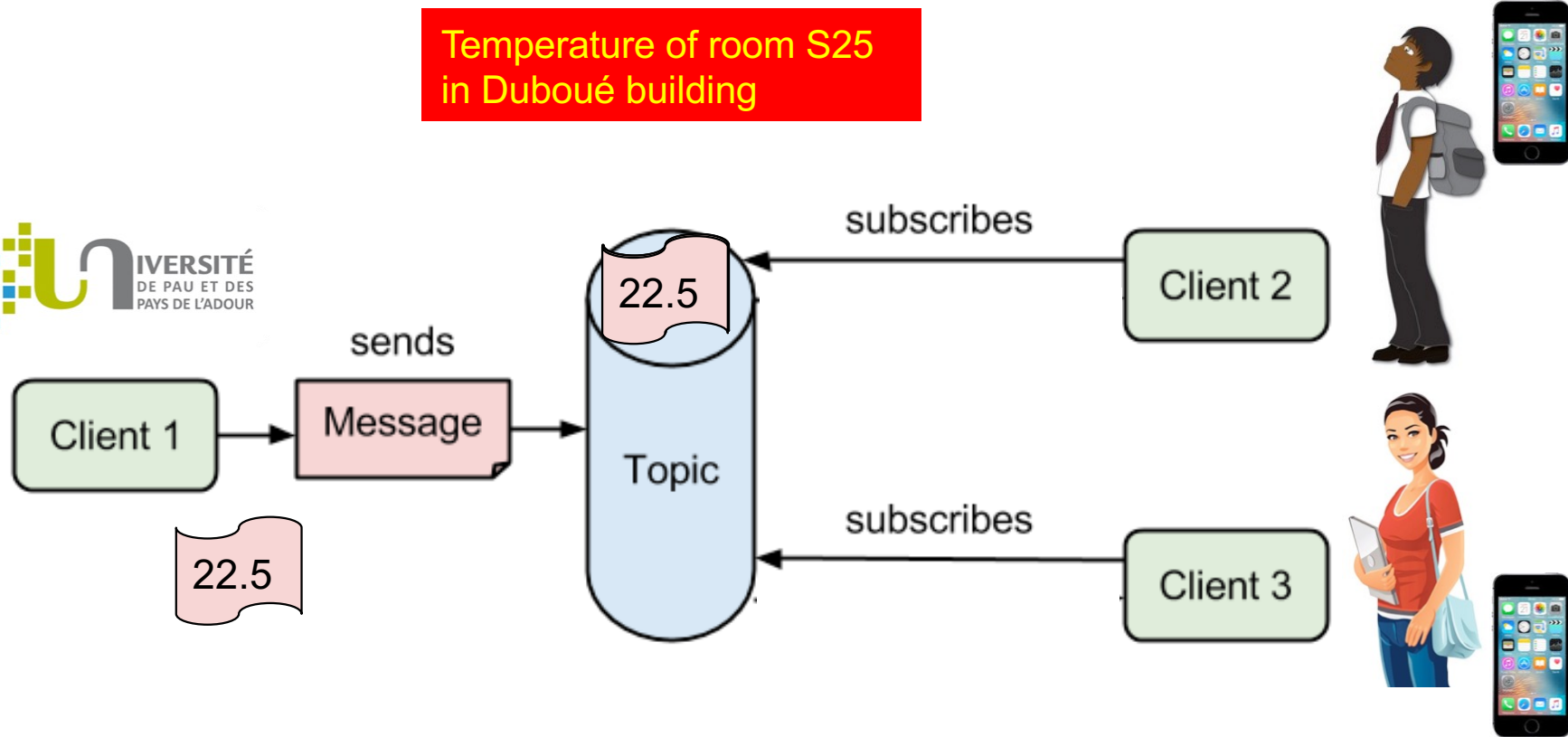
Searching for IoT data

- ⦿ Searching for information is a tough issue
 - ⦿ Web search engine: Google,...
- ⦿ If you seek for an information, for instance the soil humidity condition in a particular farm, then you need to know where to look
- ⦿ When there can be billions of IoT nodes providing large variety of data, it is difficult to find your way!
- ⦿ Although sensors' data can eventually be accessed with traditional methods (web services, HTTP/REST API, ...) IoT calls for a more "automatized" and "simplistic" approach



From "search for info" to "get the info"

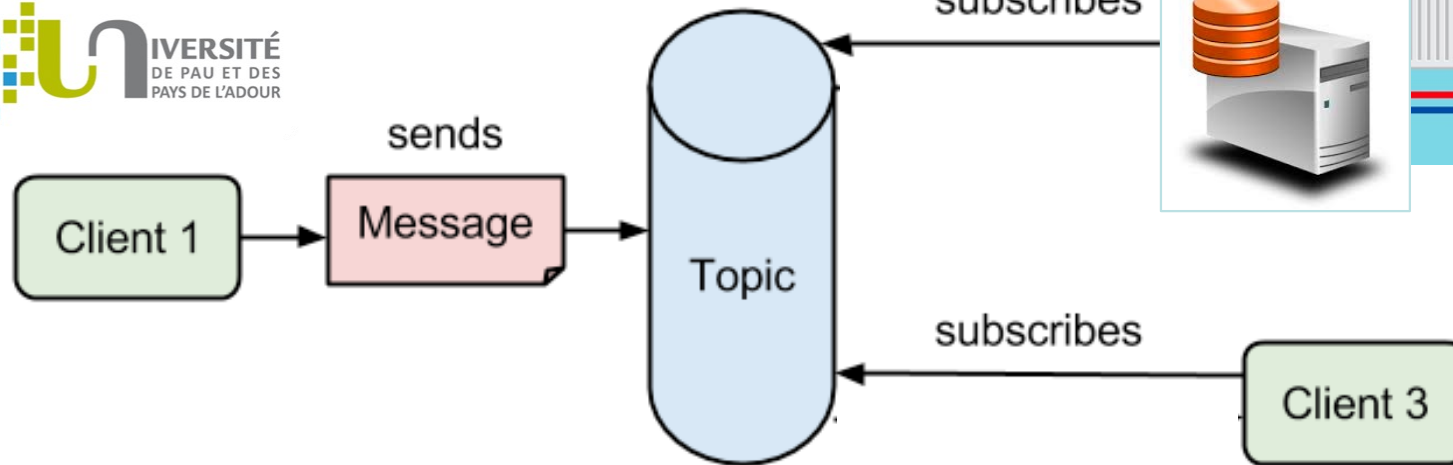
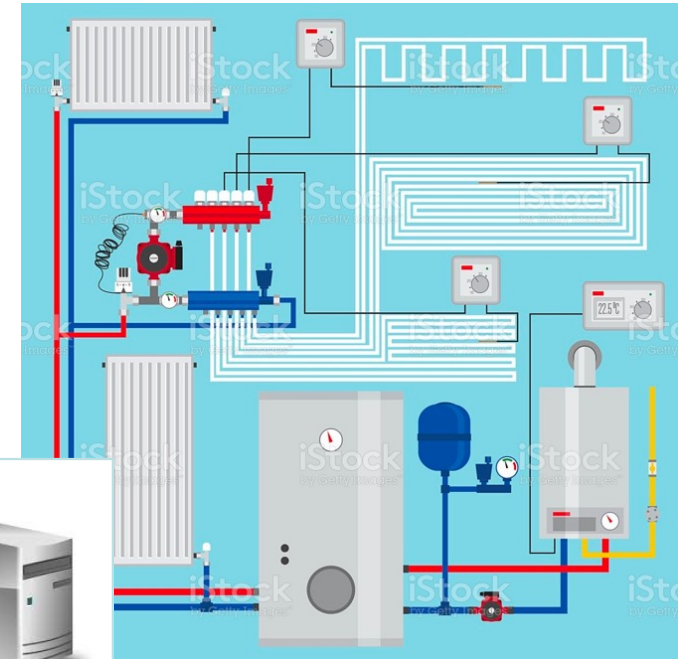
- Use the PUBLISH/SUBSCRIBE model



Automatization made simpler

- Use the PUBLISH/SUBSCRIBE model

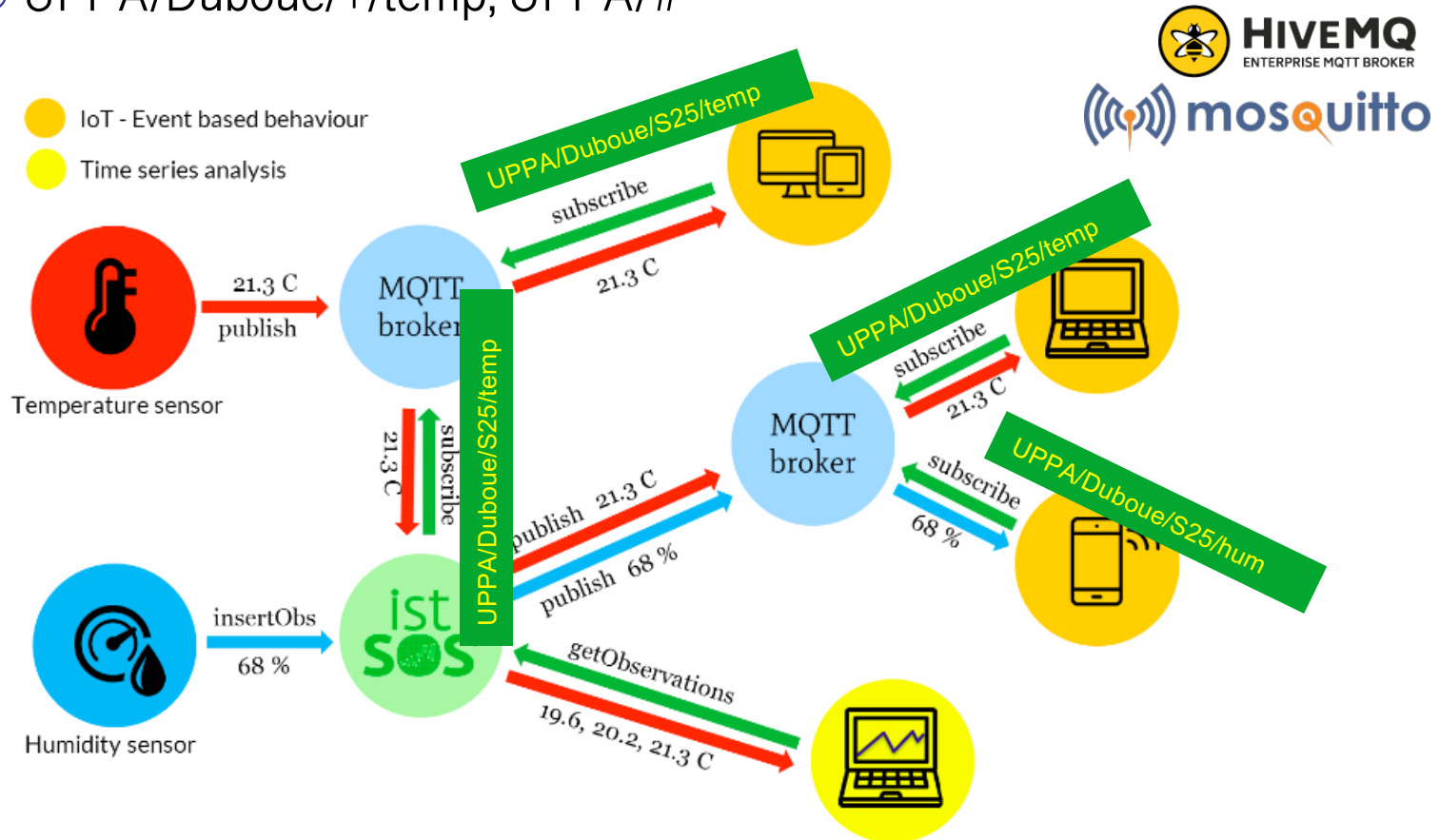
Temperature of room S25
in Duboué building



MQTT

Message Queue Telemetry Transport

- ⦿ Use broker nodes to manage topics
 - ⦿ UPPA/Duboue/S25/temp, UPPA/Duboue/S25/hum
 - ⦿ UPPA/Duboue/+/temp, UPPA/#



MQTT+smartphone=



Editors' Choice

MQTT Dash (IoT, Smart Home)

Routix software Communication ★★★★★ 1,584

PEGI 3

This app is compatible with all of your devices.

Installed

MQTT Dash

Home

My MQTT broker

Mom's house

Servers' health

My lab

MQTT Dash

This metric is intended for state displaying and switching (e.g. light on/off). Payload expected to be string.

Name

The door

Topic (sub)

door/lock

Topic (pub) - keep empty if the same as sub

Payload and icons

On 1 Off 0

Other settings

Retained

Outside humidity 11%

Garage door

Water level

☐ Towards open data

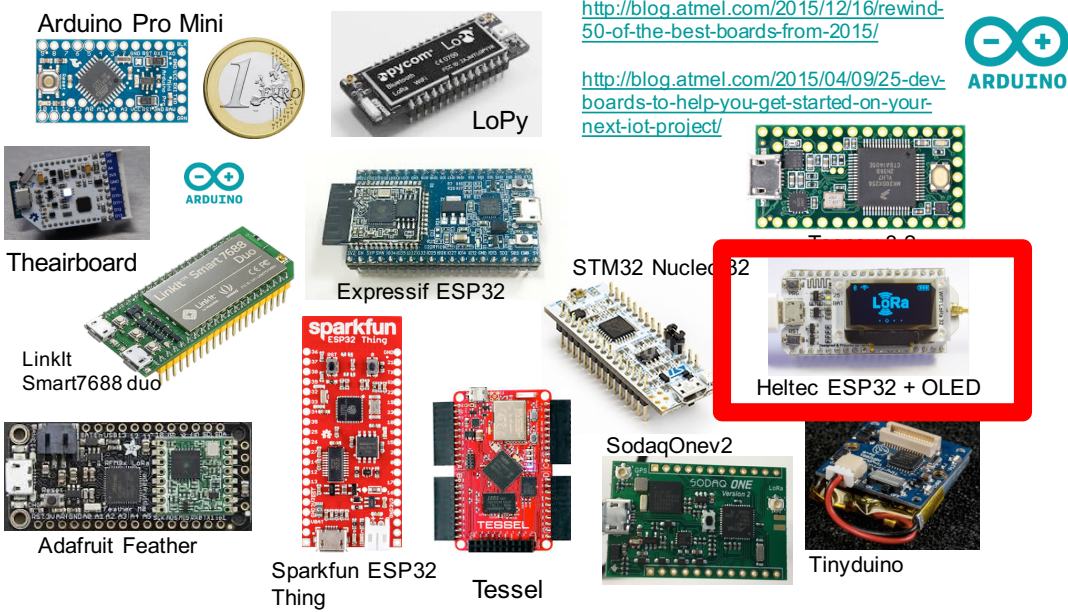
☐ UPPA/ROOMS/#

☐ UPPA/CONGRESS/#

☐ PAU/WEATHER/#

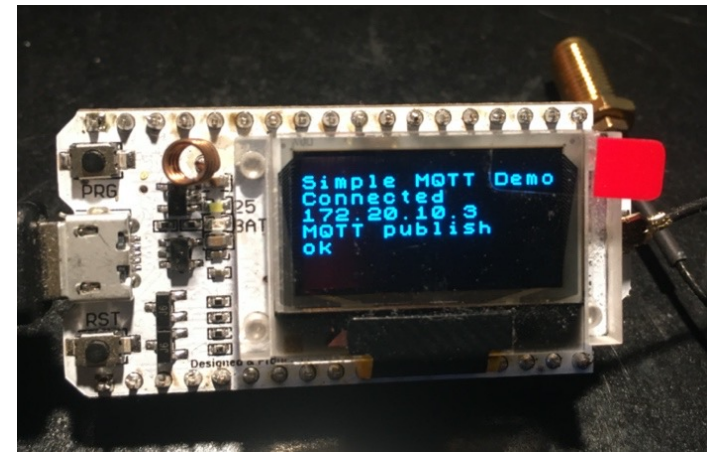
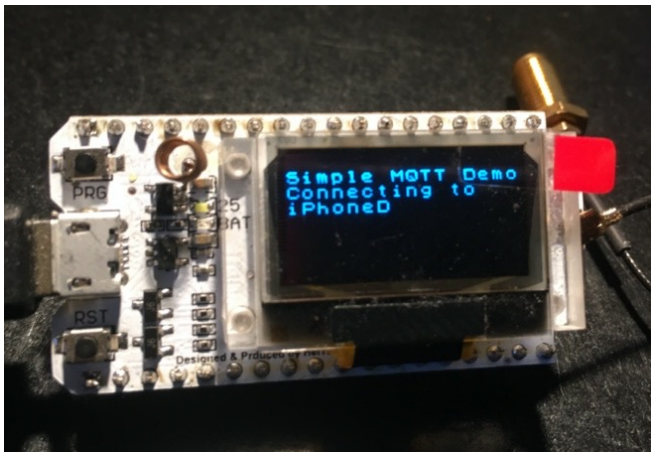
MQTT is very lightweight

- ⦿ MQTT can run on small IoT devices
- ⦿ Heltec WiFi ESP32
 - ⦿ Device connects to WiFi network
 - ⦿ Then will publish data to MQTT topic



<http://blog.atmel.com/2015/12/16/rewind-50-of-the-best-boards-from-2015/>
<http://blog.atmel.com/2015/04/09/25-dev-boards-to-help-you-get-started-on-your-next-iot-project/>

Pr. Congduc Pham
<http://www.univ-pau.fr/~cpham>



Ex: Mosquitto MQTT broker

DEMO



- ⦿ Eclipse Mosquitto is an open-source MQTT broker
- ⦿ MQTT test broker: `test.mosquitto.org`
- ⦿ IoT device will publish to topic `UPPA/Duboue/S25/temp`
- ⦿ On a computer, use `mosquitto_sub` to subscribe
 - ⦿ `mosquitto_sub -v -h test.mosquitto.org -t UPPA/Duboue/#`
 - ⦿ `-v` \Rightarrow to display information in detailed mode
 - ⦿ `-h` \Rightarrow the MQTT broker: `-h test.mosquitto.org`
 - ⦿ `-t` \Rightarrow the MQTT topic: `-t UPPA/Duboue/#`

Ex: HiveMQ broker on websocket

🕒 <http://www.hivemq.com/demos/websocket-client/>

Pr. Congduc Pham

MQTT Websocket Client

Non sécurisé | hivemq.com/demos/websocket-client/

Applications | Liste de lecture

HIVEMQ Websockets Client Showcase

Connection connected

Publish

Topic: booster_pau/test QoS: 0 Retain:

Message: hello from booster Pau

Subscriptions

Add New Topic Subscription

MQTT Websocket Client

Non sécurisé | hivemq.com/demos/websocket-client/

Applications | Liste de lecture

HIVEMQ Websockets Client Showcase

Connection connected

Publish

Topic: testtopic/1 QoS: 0 Retain:

Message:

Subscriptions

Add New Topic Subscription

Qos: 2 booster_pau/test

Messages

2021-11-25 08:55:20 Topic: booster_pau/test Qos: 0
hello from booster Pau

DEMO

MQTT in real IoT deployment

Sensor part



- Simple IoT devices have no WiFi
- Use Low-Power, Long Range radios, e.g. LoRa
- Send to IoT gateway

Control part – IoT gateway



MQTT implementing social media

- It is very easy to implement a social media app using MQTT

- WhatsApp-like example

- Define MQTT topic per phone number

- Alice: myWhatsApp/0655667788

- Bob: myWhatsApp/0611223344

- To receive/send message

- Alice publishes to myWhatsApp/0611223344

- Bob publishes to myWhatsApp/0655667788

- Both subscribe to their own topic

- To create a group

- Alice creates a group showcase-iot

- myWhatsApp/0655667788/showcase-iot

- To join(publish) on(to) the group

- Subscribe(publish) to myWhatsApp/0655667788/showcase-iot



0655667788



Alice

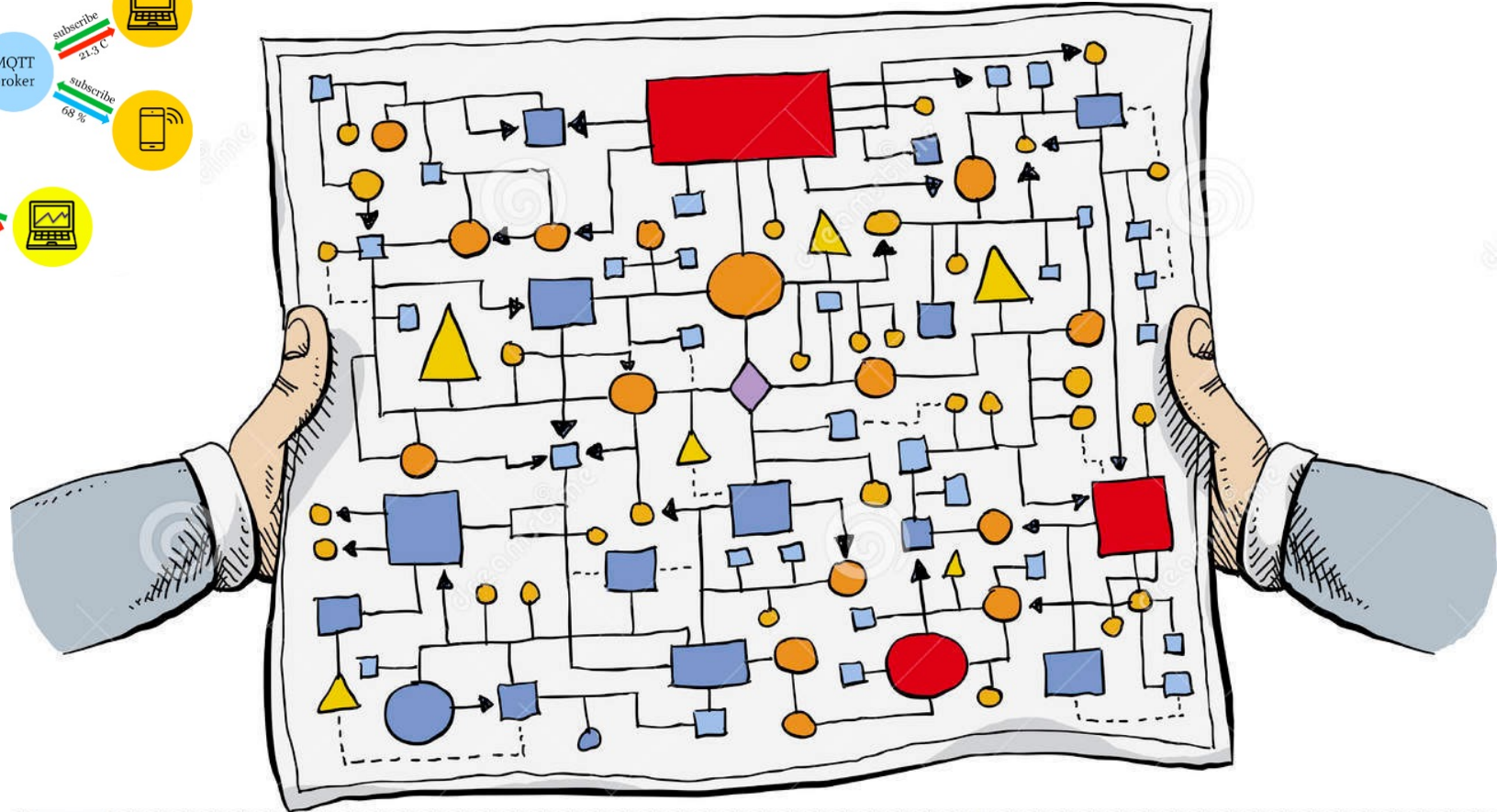
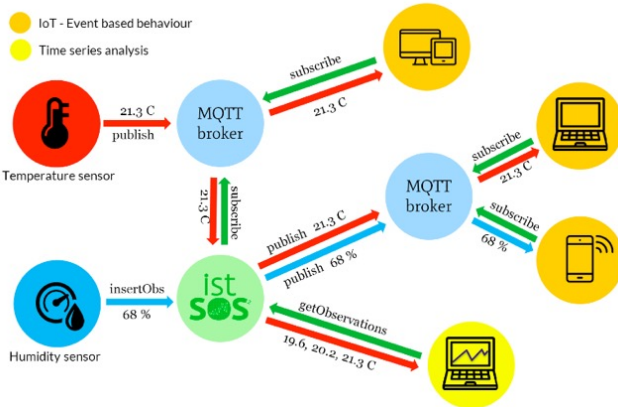


0611223344



Bob

Creating complex data flows?



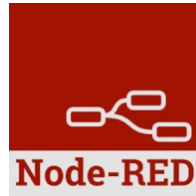
...without programming?



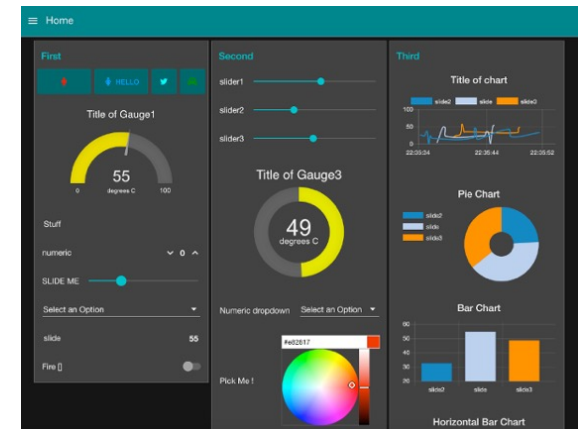
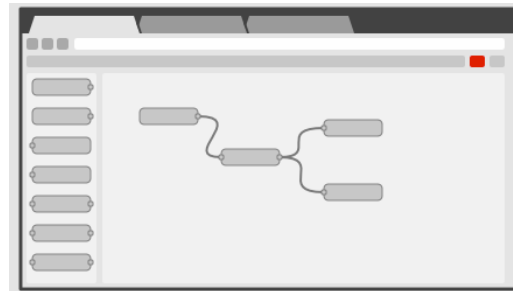
- End-users are not necessarily computer science experts nor high-skilled programmers



Node-RED

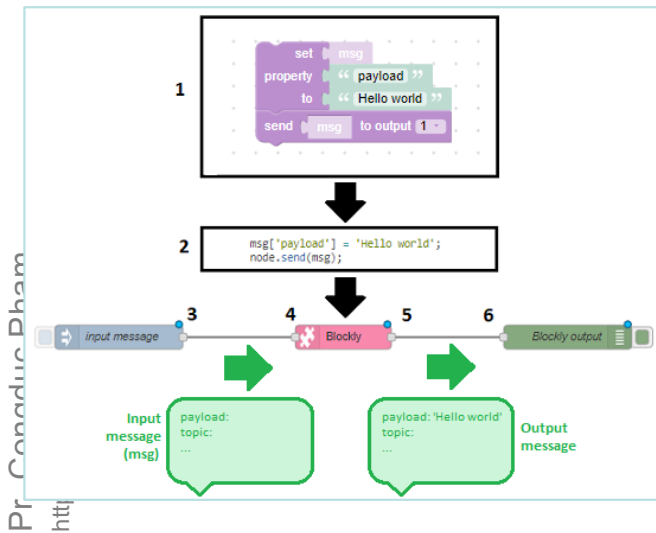


- Use graphical tools to build data processing flows, allowing intuitive connection from IoT data producers to IoT data consumers
- Node-RED is a programming tool for wiring together hardware devices, APIs and online services, e.g. clouds of various types
- provides a browser-based flow editor to wire together flows with a wide range of nodes



Node-RED blocks

Increasing number of Node-RED blocks



Nodes represent reusable pieces of code and logic. Node-RED comes with a core set of useful nodes, but there are a growing number of additional nodes available to install from both the Node-RED project as well as the wider community or you

Nodes

input

- inject
- catch
- status
- mqtt
- http
- websocket
- serial
- tcp
- mqtt
- ibmiot

output

- debug
- mqtt
- http response
- websocket
- serial
- tcp
- udp
- mqtt
- twilio
- ibmpush
- ibmiot
- OpenWhisk

function

- function
- template
- delay
- trigger
- comment
- http request
- switch
- change
- range
- csv
- html
- json
- xml
- rbe
- tcp request
- OpenWhisk

social

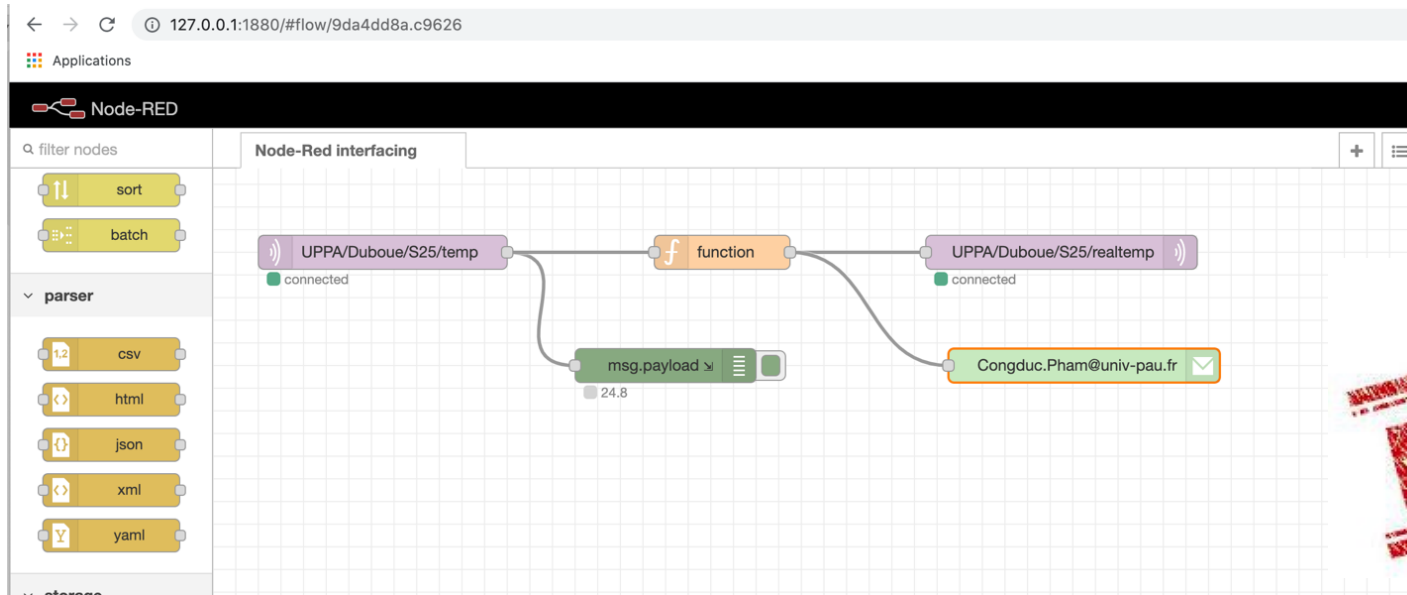
- e mail
- twitter
- e mail
- twitter

storage

- mongodb
- ibm hdfs
- ibm hdfs
- cloudant
- sqldb
- dashDB
- mongodb
- cloudant
- sqldb
- dashDB

Simple MQTT Node-RED flow

- ⦿ MQTT in-broker: `test.mosquitto.org`
- ⦿ **"MQTT in"** node listens on `UPPA/Duboue/S25/temp`
- ⦿ "Function" node to correct temperature by -1.8°C
- ⦿ MQTT out-broker: `broker.hivemq.com`
- ⦿ **"MQTT out"** node publishes on `UPPA/Duboue/S25/realtemp`
- ⦿ "Mail" node sends corrected temp to `Congduc.Pham@univ-pau.fr`



DEMO

Simple MQTT Node-RED flow

The image shows a Node-RED MQTT flow on the left and the Hivemq Websocket Client interface on the right. The Node-RED interface includes a sidebar with nodes like 'sort', 'batch', 'parser' (with 'csv', 'html', 'json', 'xml', 'yaml' sub-nodes), and 'storage'. The MQTT flow consists of a 'mqtt in' node connected to a 'mqtt out' node. The Hivemq interface shows a 'connected' status, a 'Publish' section with 'testtopic/1' as the topic, and a 'Subscriptions' section with two active subscriptions: 'booster_pau/test' and 'UPPA/Duboue/S25...'. The 'Messages' section displays a list of received messages:

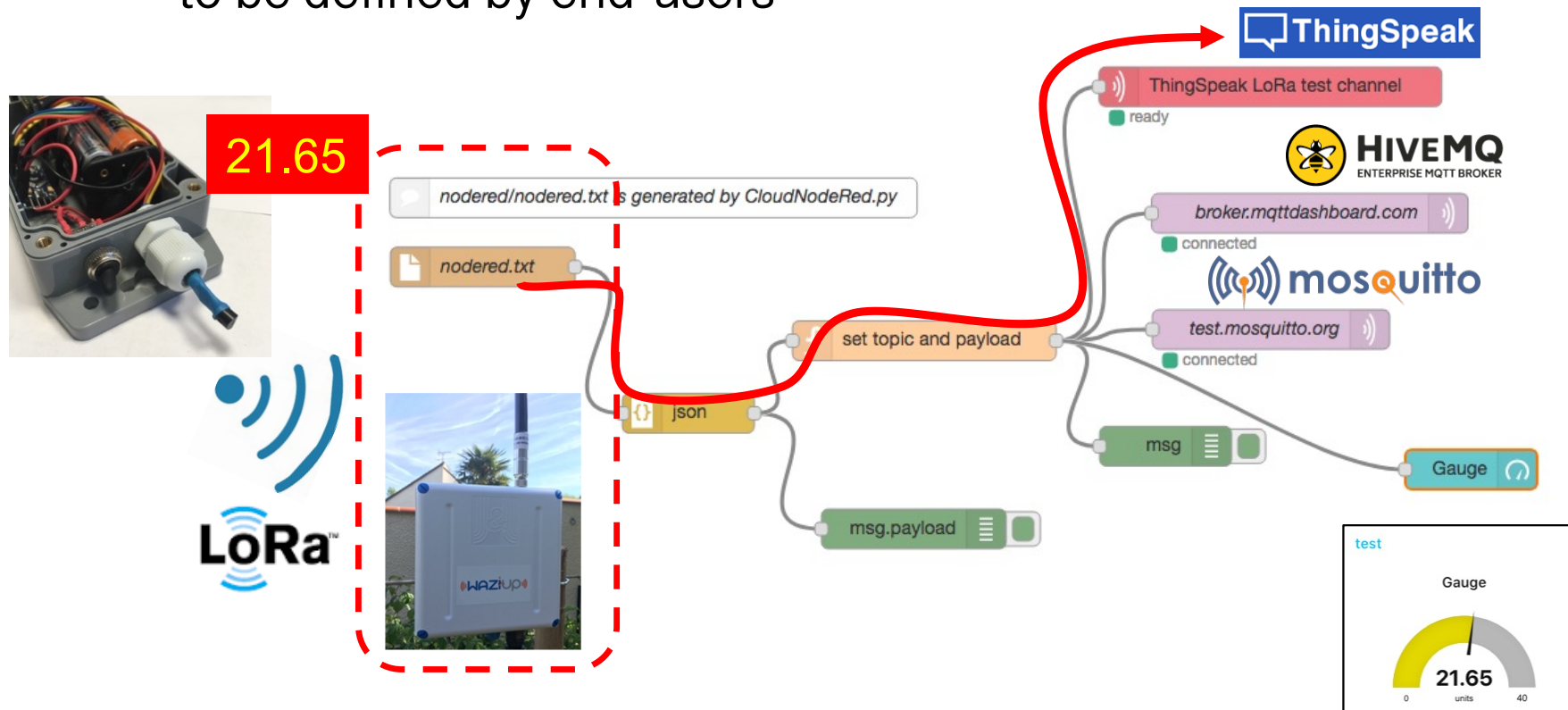
Timestamp	Topic	Message	QoS
2021-11-25 09:18:14	UPPA/Duboue/S25/realtemp	20.7	0
2021-11-25 09:18:11	booster_pau/test	hello from booster Pau	0
2021-11-25 09:18:04	UPPA/Duboue/S25/realtemp	20.7	0
2021-11-25 09:17:57	UPPA/Duboue/S25/realtemp	20.7	0

Pr. Congduc Pham
<http://www.univ-pau.fr/~cpham>

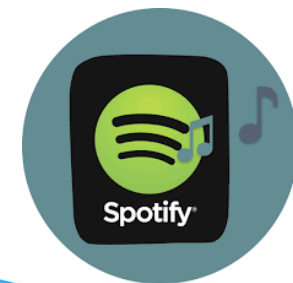
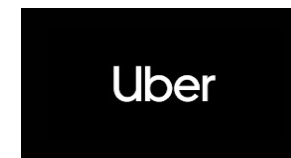


Node-RED enabled IoT gateway

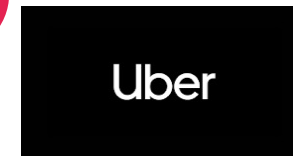
- Messages received on the IoT gateway can be injected into a Node-RED flow, allowing complex data processing to be defined by end-users



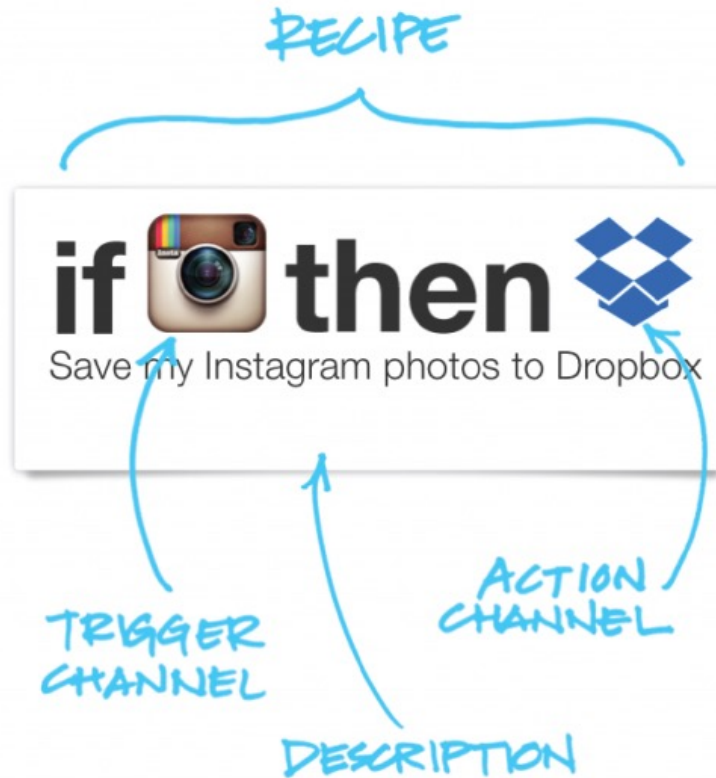
Generalizing interactions?



Adding interactions?



IF THIS THEN THAT applets



Some example Recipes

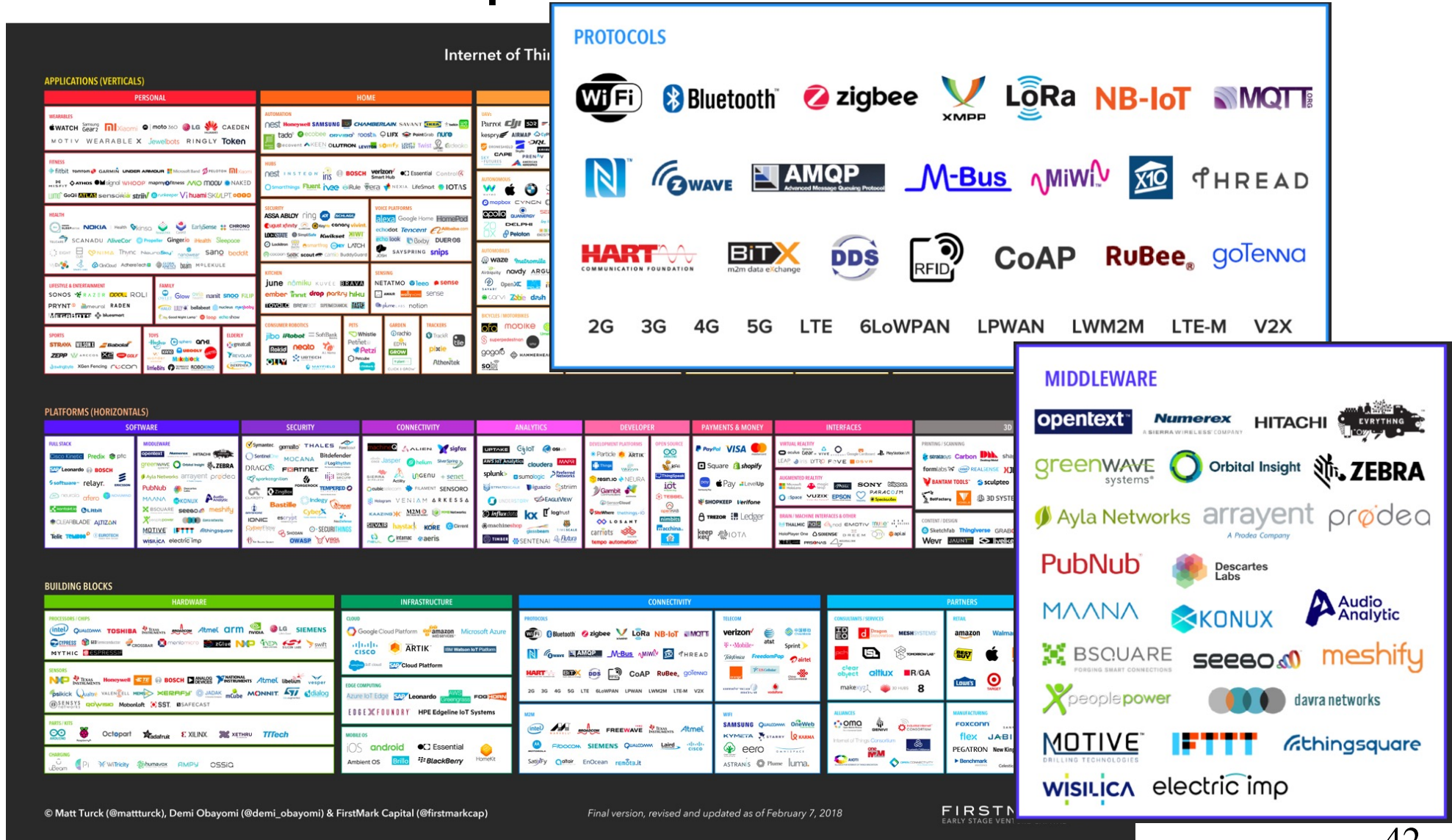
if  **then** 
Nearly home? Direct message the person who should know

if  **then** 
Email your new iPhone photos to yourself

if  **then** 
Backup your contacts to a Google Spreadsheet

IoT landscape

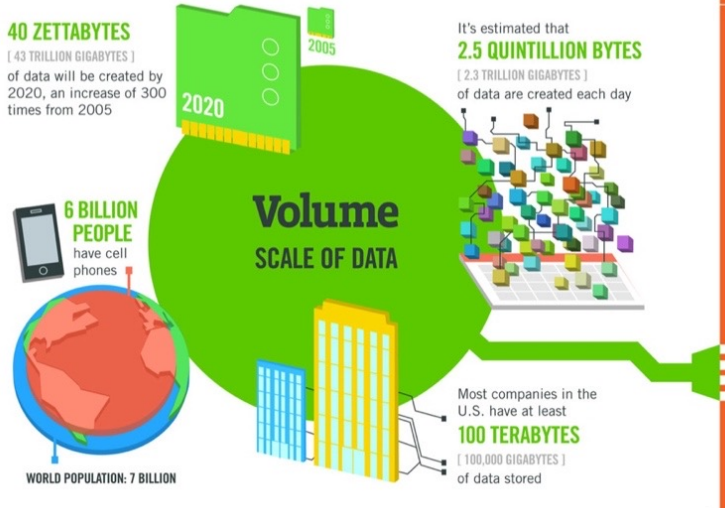
Pr. Congduc Pham
http://www.univ-pau.fr/~cpham





IOT BACKOFFICE

IoT usually means **BIG DATA** ...



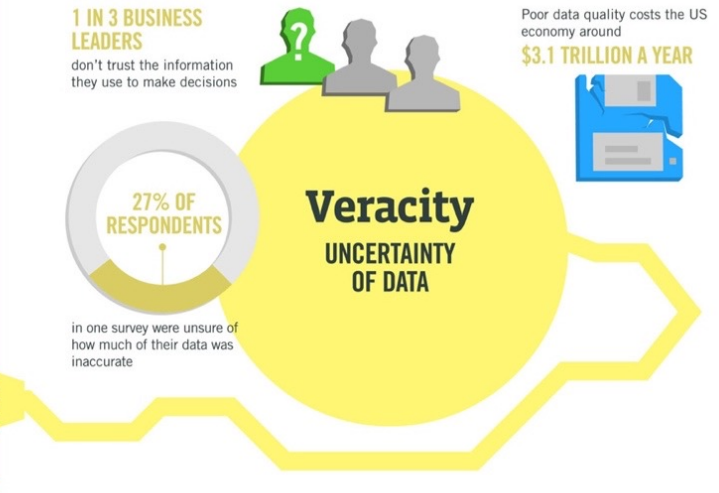
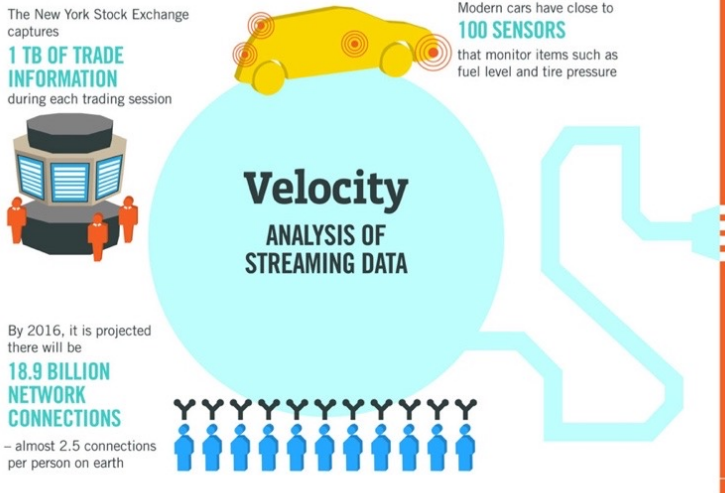
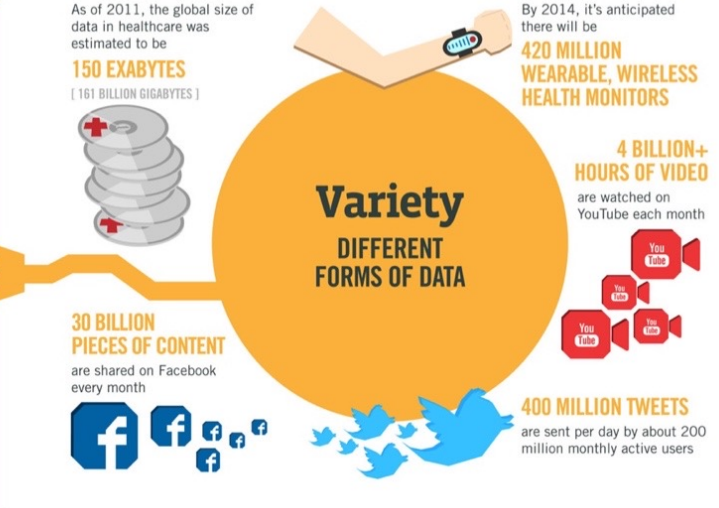
The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

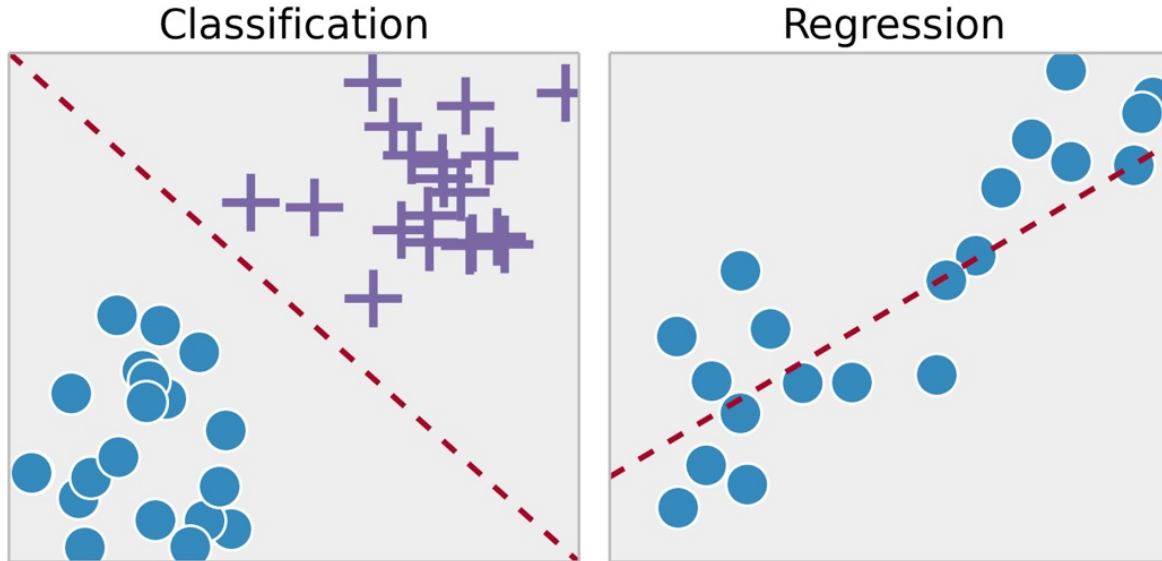
Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015 **4.4 MILLION IT JOBS** will be created globally to support big data, with 1.9 million in the United States



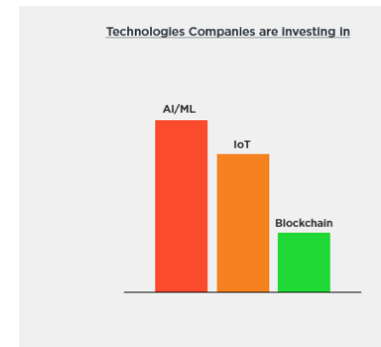
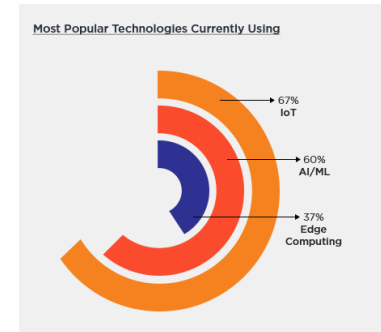
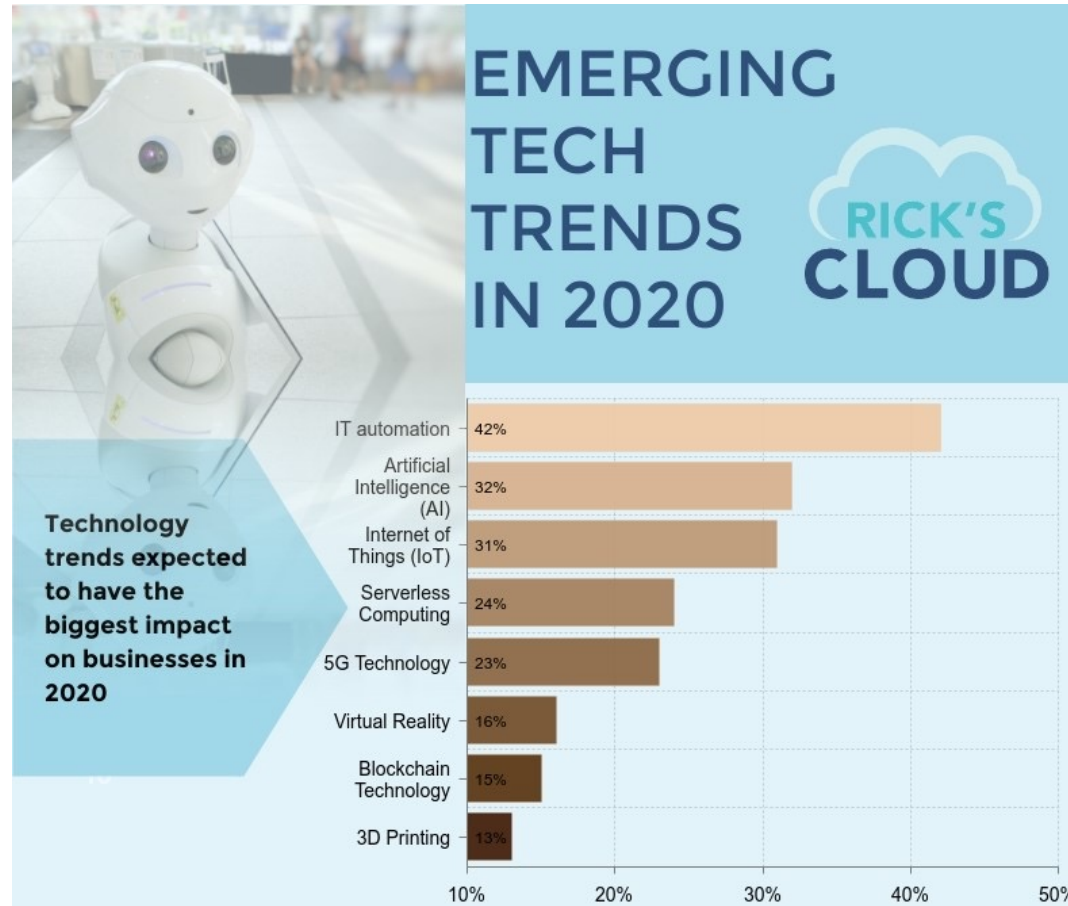
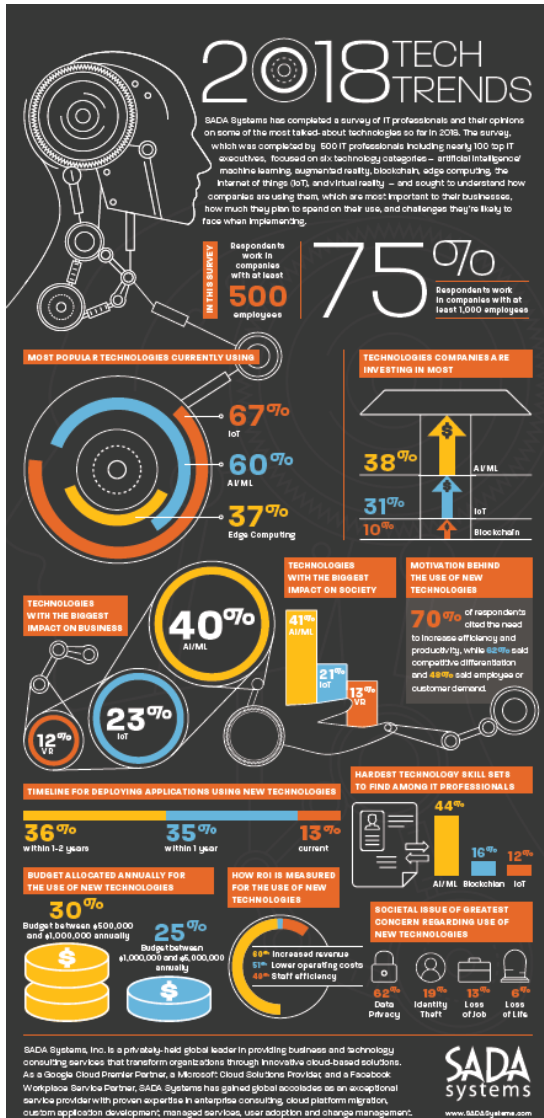
...but also how to analyse the data

- ⦿ What is the meaning of the collected data?
 - ⦿ Classification
- ⦿ Can we predict how the data will evolve?
 - ⦿ Prediction using Regression methods



Analysing IoT data: what's the trend?

Pr. Congduc Pham
http://www.univ-pau.fr/~cpham

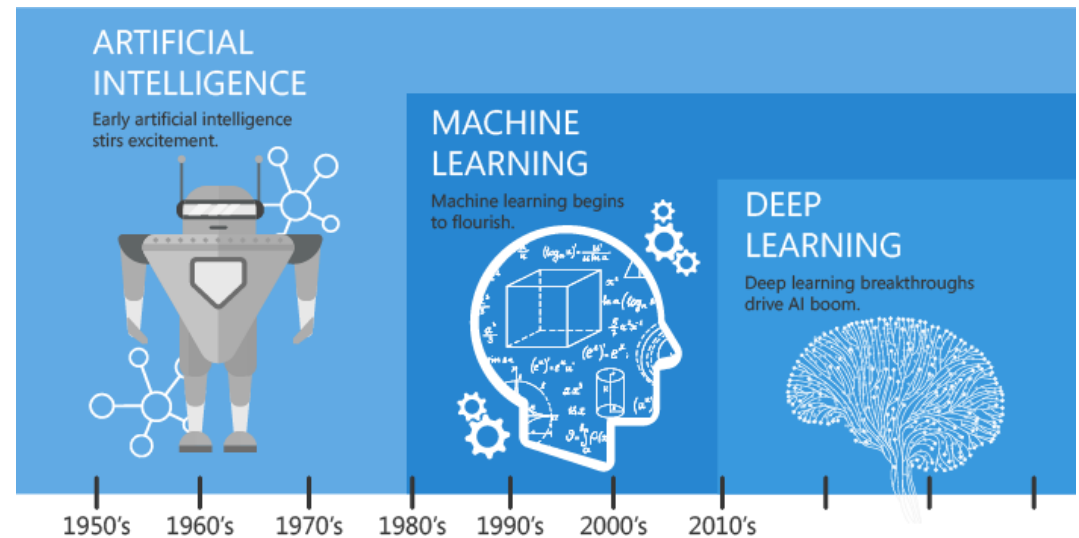


There must be a reason IoT+AI are number 1!

The raise of Artificial Intelligence

- It is the science and engineering of making intelligent machines.
- In Computer Science, Artificial Intelligence (**AI**) research is defined as the study of « intelligent agents »
- From General AI to Narrow AI: from overhyping to fewer promises, but more realistic!

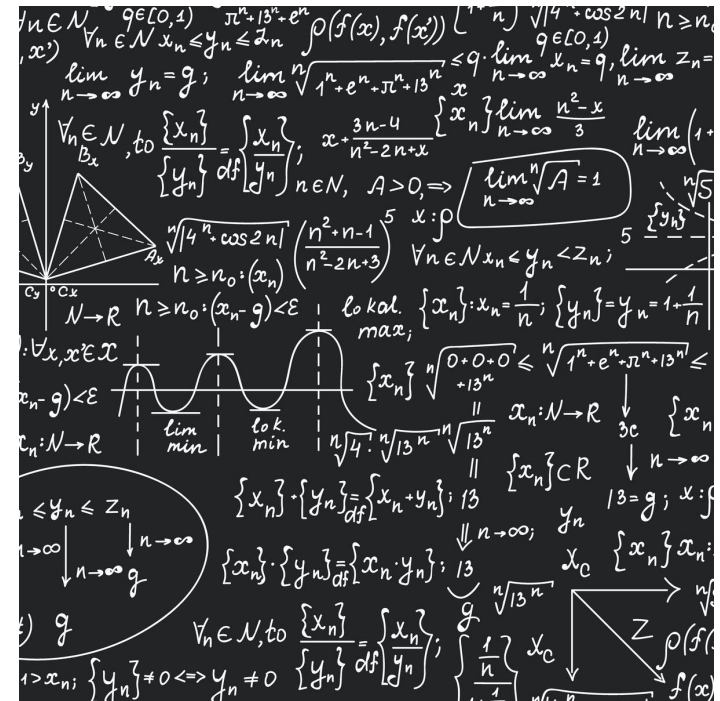
Pr. Congduc Pham
<http://www.univ-pau.fr/~cpham>



Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

AI: (now) a serious science!

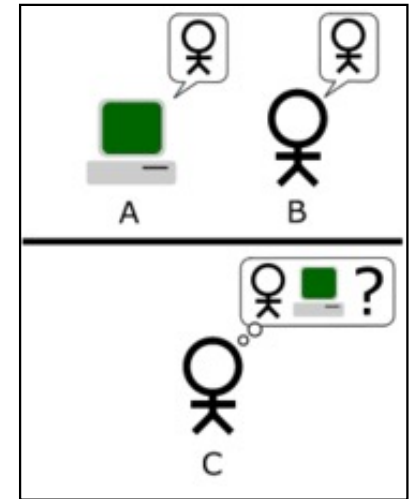
- ⦿ General-purpose AI like the robots of science fiction is incredibly hard
 - ⦿ Human brain appears to have lots of special and general functions, integrated in some amazing way that we really do not understand (yet)
- ⦿ Special-purpose AI is more doable (nontrivial)
 - ⦿ E.g., chess/poker playing programs, logistics planning, automated translation, speech and image recognition, web search, data mining, medical diagnosis, keeping a car on the road.



The Turing Test



- Proposed By Alan Turing in 1950
- To be called intelligent, a machine must produce responses that are indistinguishable from those of a human.
- Human judge communicates with a human and a machine over text-only channel.
- Both human and machine try to act like a human. Judge tries to tell which is which.
- Is Turing Test the right goal?



“Aeronautical engineering texts do not define the goal of their field as making ‘machines that fly so exactly like pigeons that they can fool even other pigeons.’”
[Russell and Norvig]

Reflection

if AI can be more
rational than
humans in some
cases, why not?



Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that <u>act</u> <u>rationally</u>

AI focus on **action**.
Avoids philosophical
issues such as “is the
system conscious” etc.

AI Technologies

AI TECHNOLOGIES

STRONG / GENERAL AI



WEAK / SPECIALIZED KI

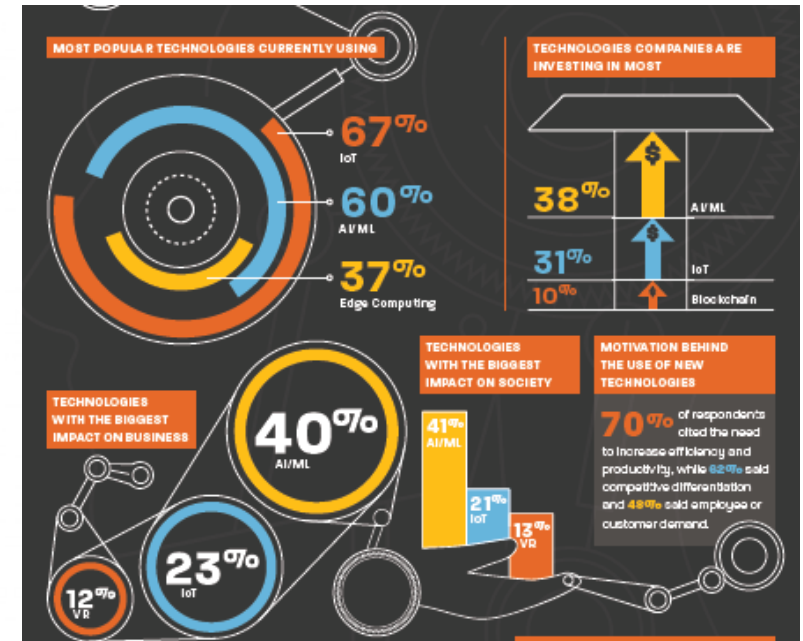
CLASSIC / SYMBOLIC

DECISION TREES	TABLE BASED AGENTS
SEARCH	EXPERT SYSTEMS
RULE BASED SYSTEMS	SYMBOLIC LOGIC

EXPERT SYSTEMS WITH NN INPUT	DECISION TREE LEARNING
MONTE CARLO SEARCH WITH NN	XGBOOST
	MIXED ALGORITHMS

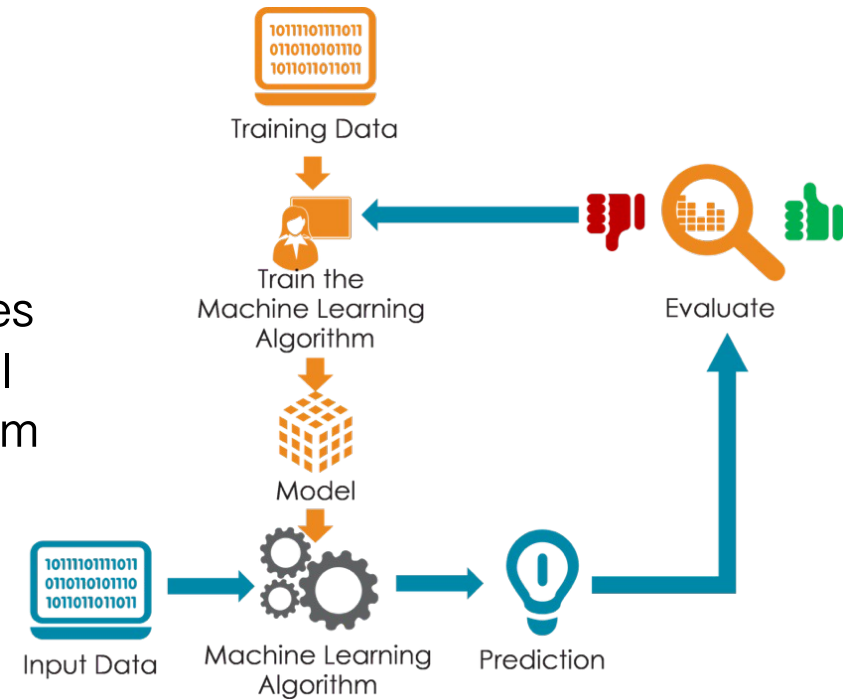
NEURAL NETWORKS	LINEAR REGRESSION
LSTM	NAIVE BAYES
CONVOLUTIONAL NN	RANDOM FOREST
ATTENTION	Q LEARNING
AUTOENCODER	EVOLUTIONARY ALGORITHMS
DEEP LEARNING	

MACHINE LEARNING



Machine Learning

- Develops Narrow Artificial Intelligence systems through examples
 - A developer creates a model and then “trains” it by providing it with many examples
 - The machine learning algorithm processes the examples and creates a mathematical representation of the data that can perform prediction and classification tasks
- Example
 - A machine-learning algorithm trained on thousands of bank transactions with their outcome (legitimate or fraudulent) will be able to predict if a new bank transaction is fraudulent or not



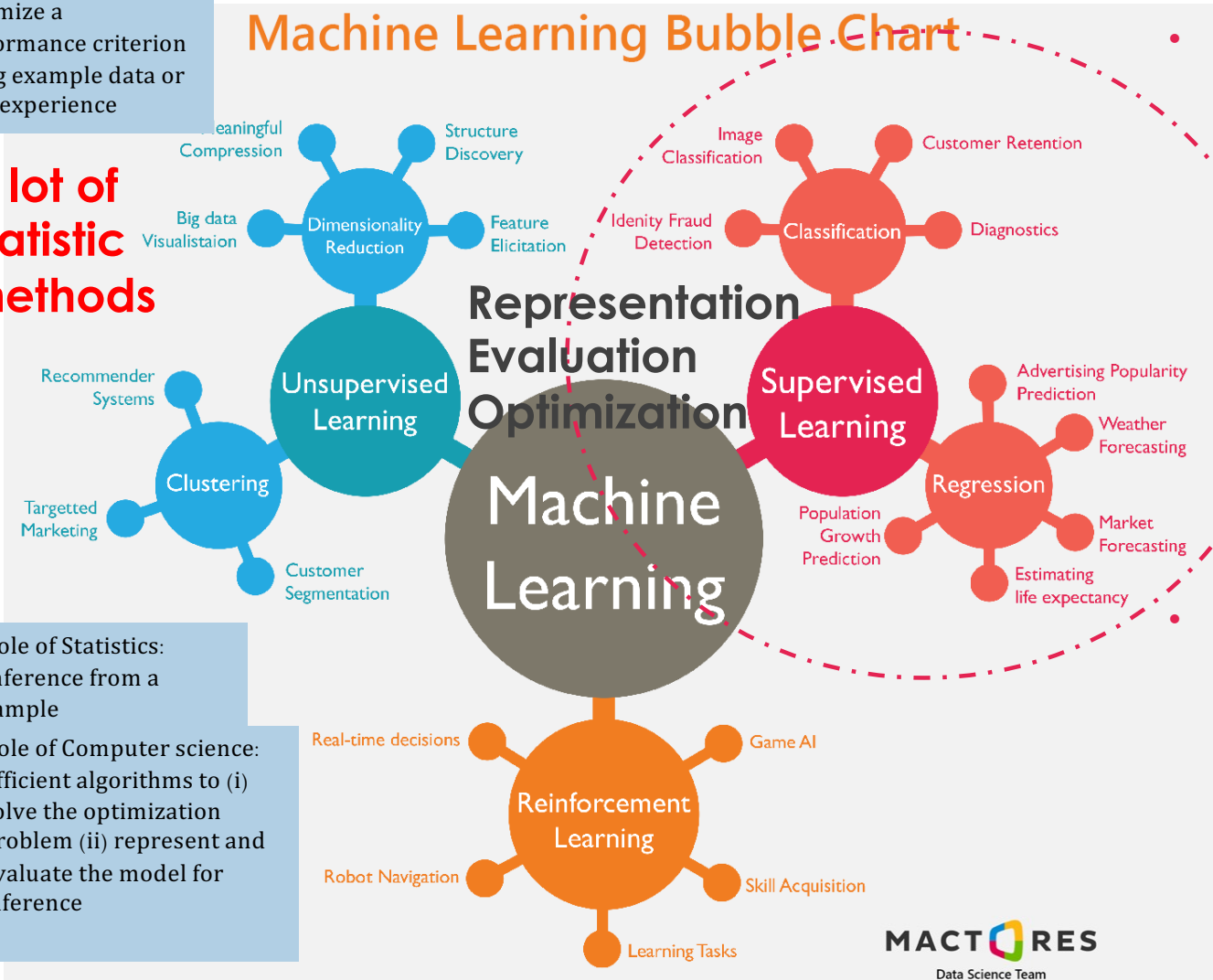
Machine Learning Techniques

Optimize a performance criterion using example data or past experience

Machine Learning Bubble Chart

A lot of statistic methods

Role of Statistics: Inference from a sample
 Role of Computer science: Efficient algorithms to (i) solve the optimization problem (ii) represent and evaluate the model for inference



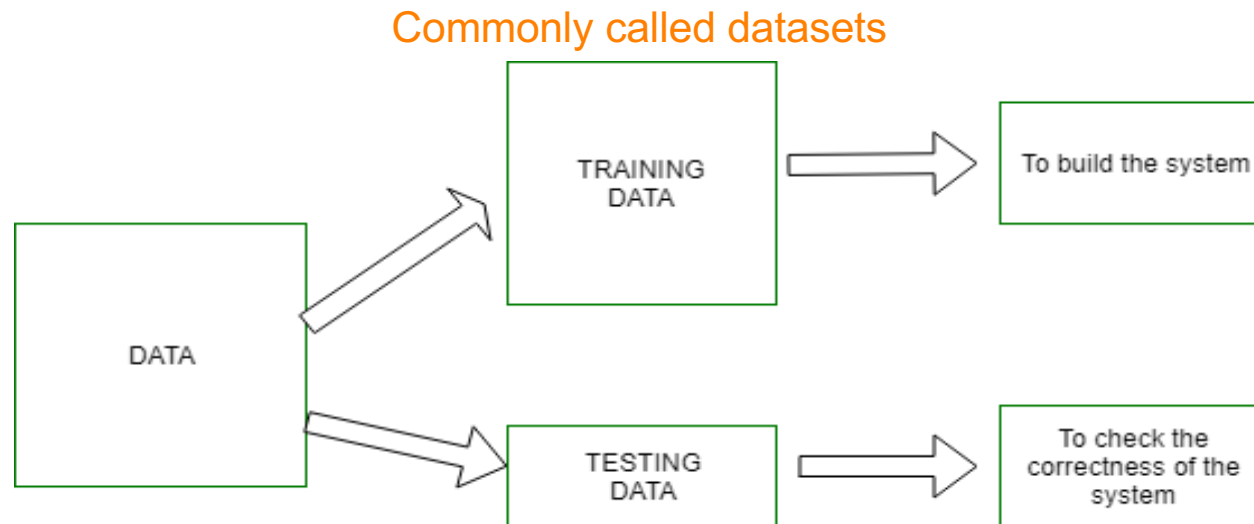
- Classification
 - Logic
 - SVM
 - Random Forest
 - Hidden Markov

- Regression
 - Lasso
 - Ridge
 - Loes
 - KNN
 - Spline
 - XGBoost

Pr. Congduc Pham
<http://www.univ-pau.fr/~cpham>

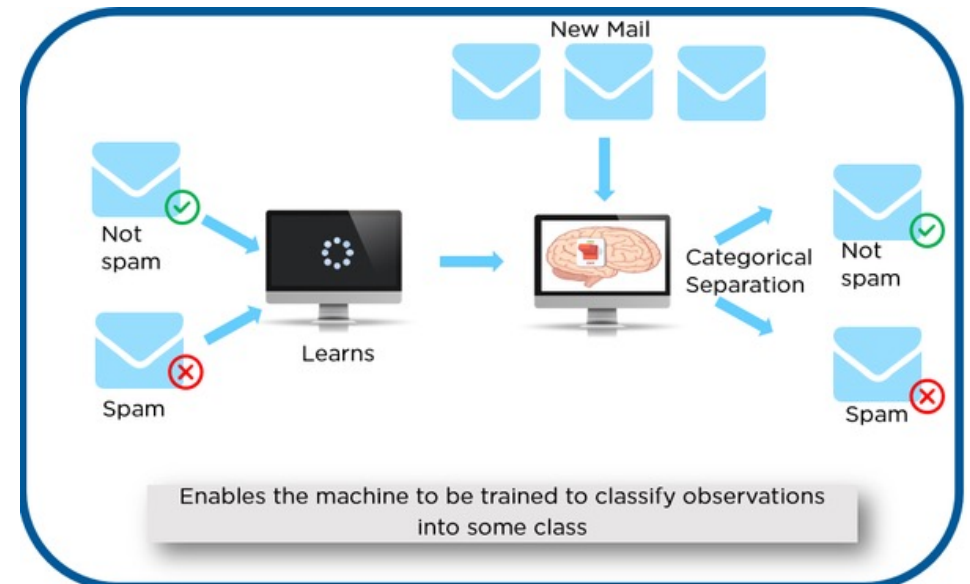
Supervised Learning

- ⦿ ML model is presented with *input data* which **is labeled**
 - ⦿ Each *input data* is tagged with the correct label
- ⦿ The goal is to approximate math operations in the ML model so well that when presented with new *input data*, the ML model can **predict** the output variables for that *input data*.



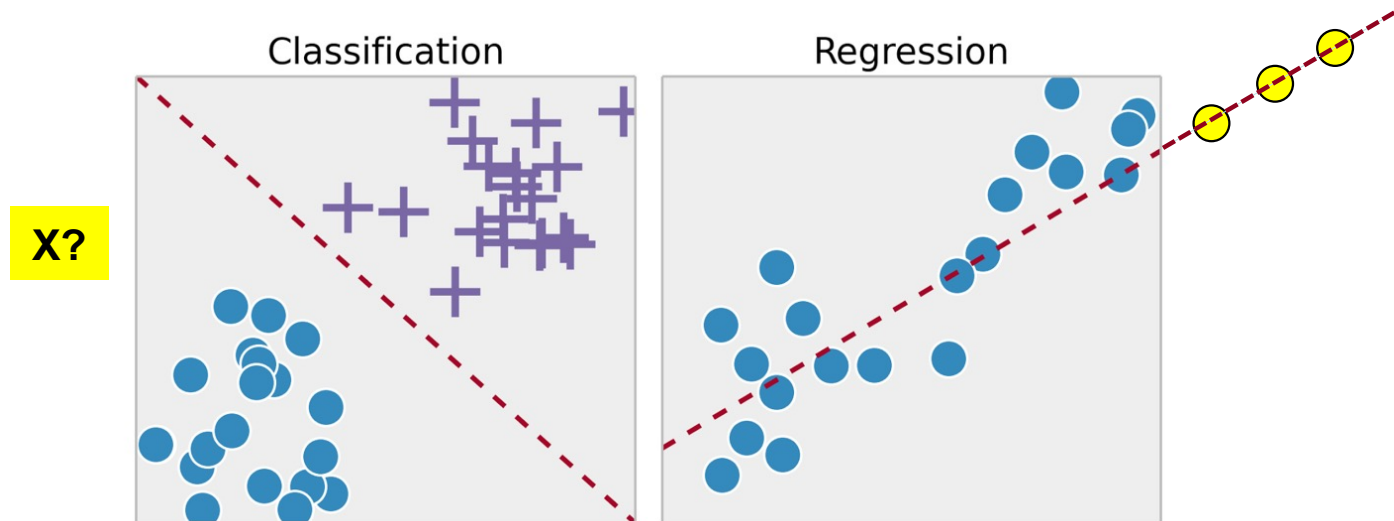
Spam Mail Example

- On the left side of the image, some data is marked as ‘Spam’ or ‘Not Spam’. This is *labeled data*. This data is used to train the supervised model, the *intelligent* program (at center of the image).
- Trained model is tested with new mails (on the top of the image) and checking if the output of the supervised model is correct (on the right side of the image).



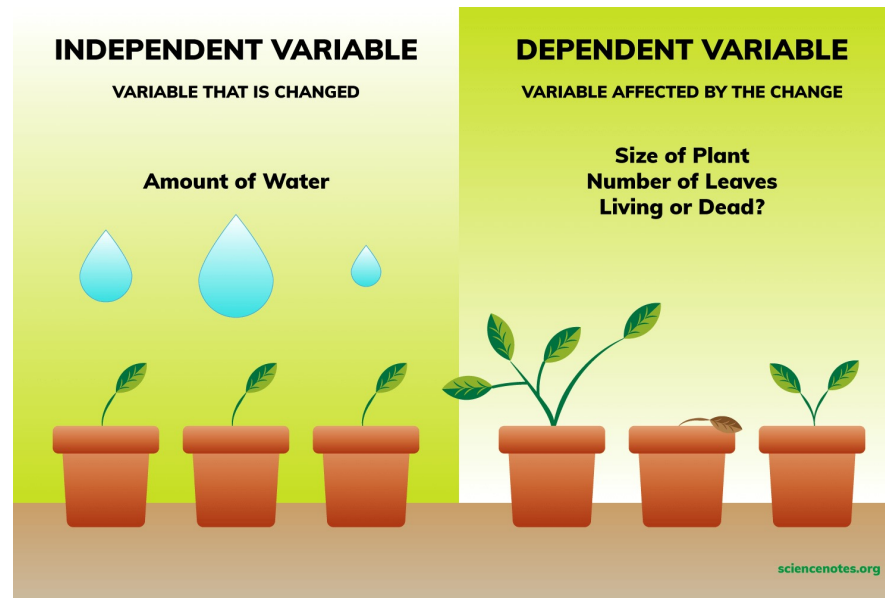
Types of Supervised Learning

- ⦿ **Classification:** A classification problem is when the output is a category, such as “red” or “blue” or “disease” and “no disease”.
- ⦿ **Regression:** A regression problem is when the output is a real number, such as “dollars” or “weight”.



Regression

- ⦿ **Dependent variables:** the main event or factor to understand or predict. Also known as *explanatory variable*.
- ⦿ **Independent variables:** the events or factors suspected to have an impact on the dependent variable. Also known as *response variable*.



Types of Regression

- ⦿ **Simple regression:** single independent variable x for a single dependent variable Y .

x : number of cricket chirps

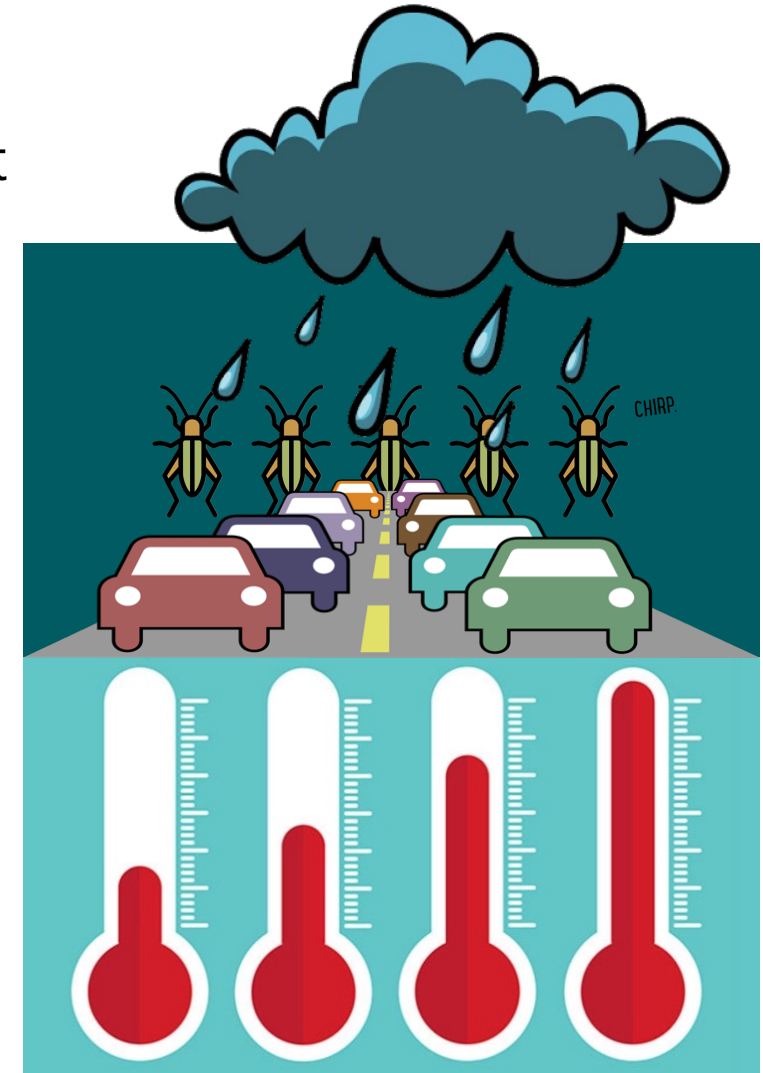
Y : temperature
- ⦿ **Multivariable regression:** multiple independent variables, x_1, x_2, x_3 , for a dependent variable Y .

x_1 : number of cricket chirps

x_2 : rainfall

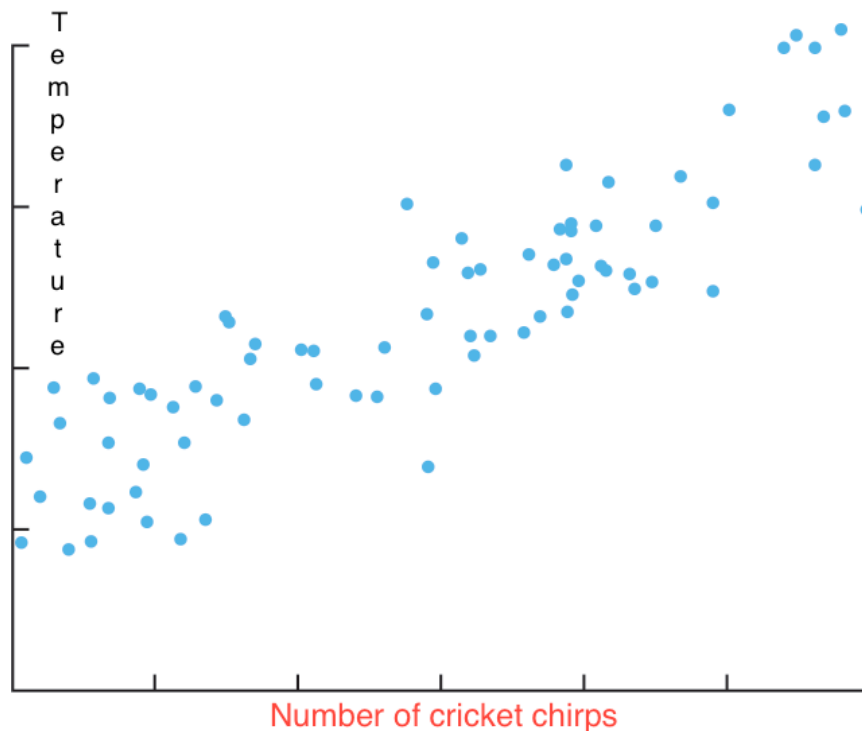
x_3 : automobile traffic

Y : temperature



Scatter Plot of datasets

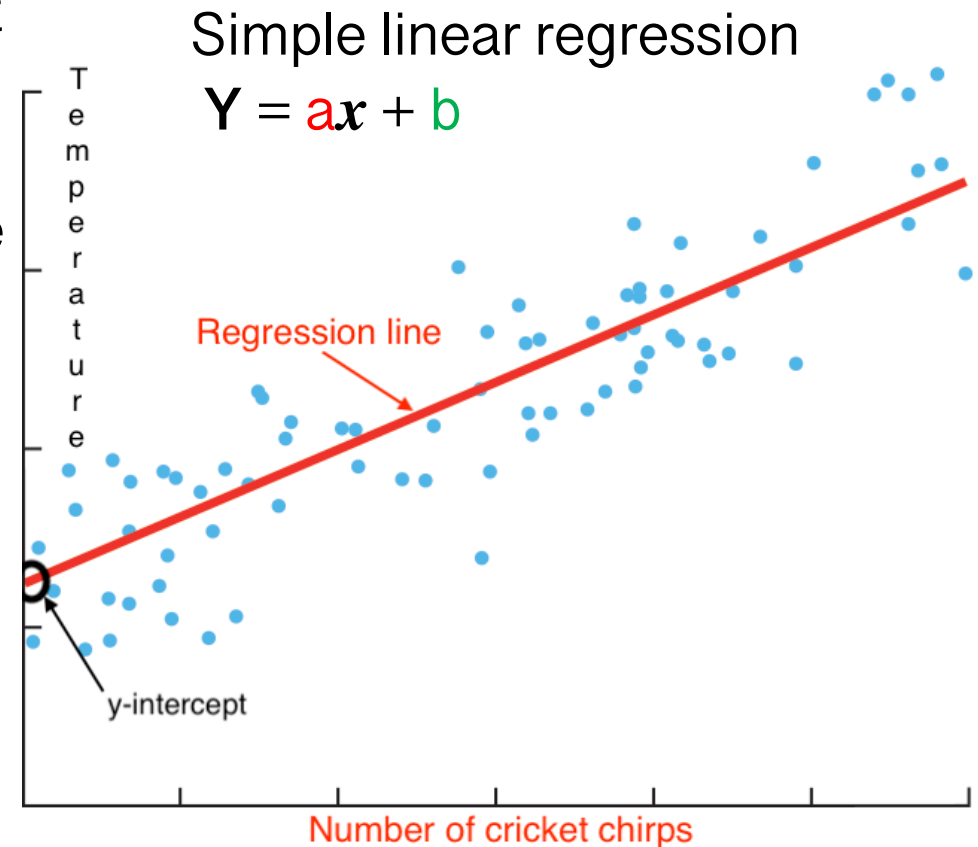
- ⊙ Data gathering on the variables in question
- ⊙ The vertical scale represents one set of measurements and the horizontal scale the other



- ⊙ Agriculture
 - ⊙ Soil moisture for optimized irrigation
- ⊙ Environment
 - ⊙ Pollutant concentration for public safety and alarming system
- ⊙ Electricity
 - ⊙ Power demands for optimized production
- ⊙ ...

Linear Regression

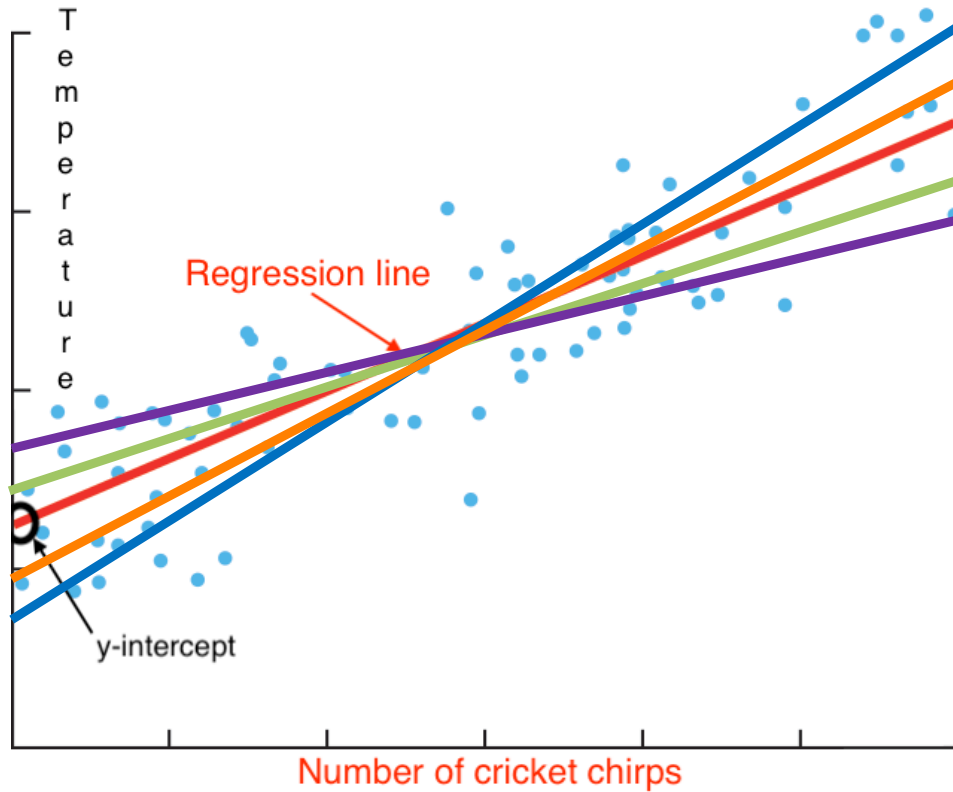
- ⦿ A linear relationship to predict the (average) numerical value of Y for a given value of x using a straight line, called the *regression line*.
- ⦿ Knowing the *slope* and the *y*-intercept the objective is to predict the average Y from x .
- ⦿ Real world problem can be multivariate
- ⦿ Multiple linear regression



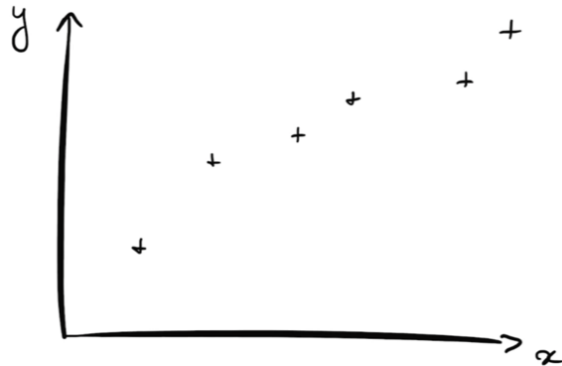
$$Y = a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_ix_i + b$$

Then what?

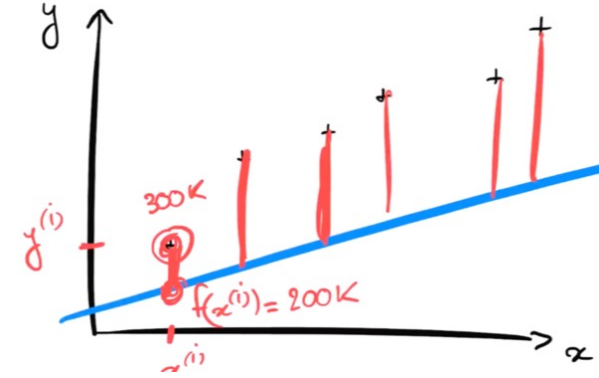
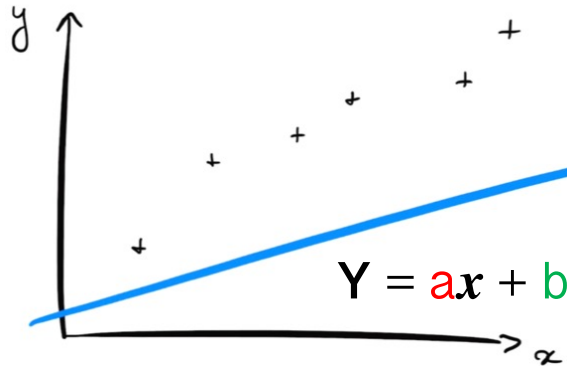
- ⦿ What is the best regression line?
- ⦿ How can we find it?



The need for a cost function

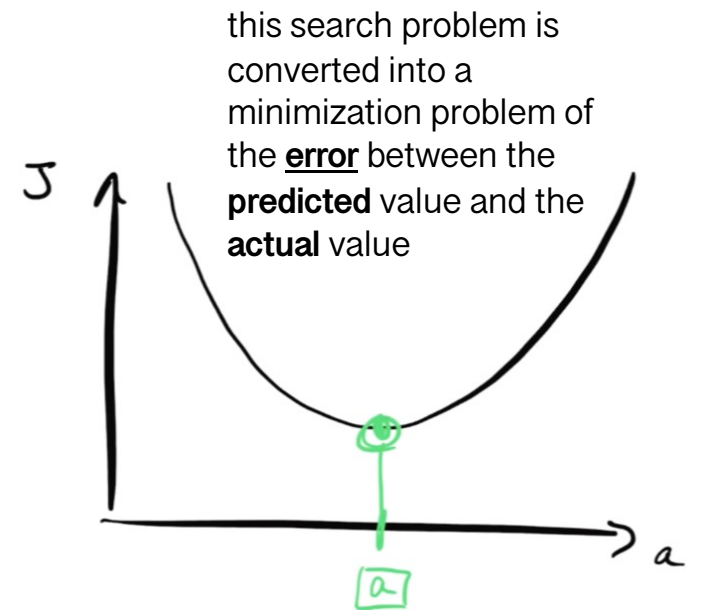
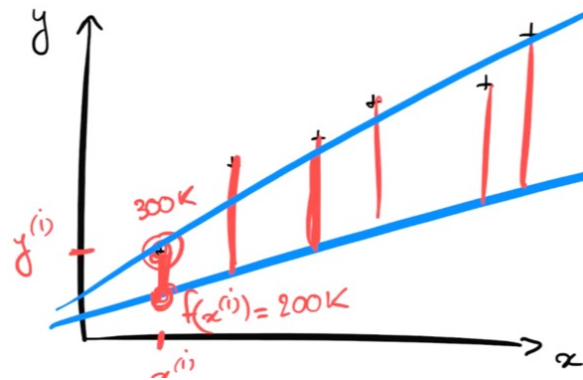


1. Dataset: (x, y) $m = 6, n = 1$ $x^{(i)}$



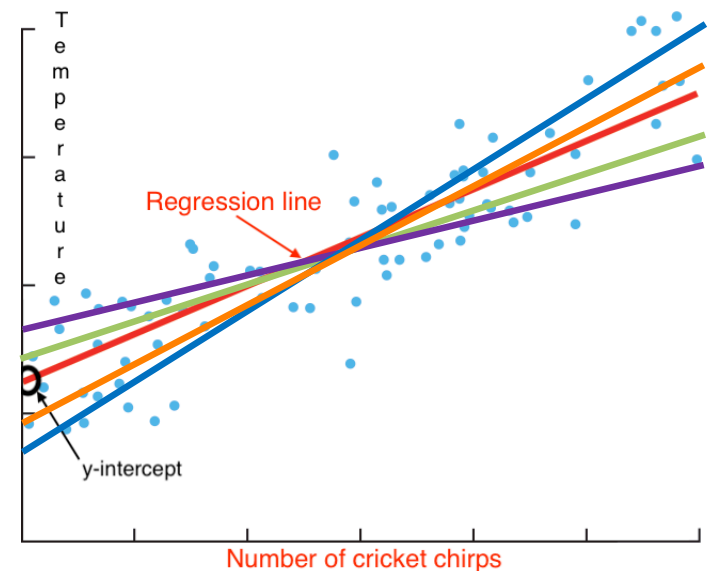
$$J(a, b) = \frac{1}{2m} \sum_{i=1}^m (f(x^{(i)}) - y^{(i)})^2$$

A **cost function** will help to figure out the best possible values for **a** and **b** which would provide the best *regression line* for data points



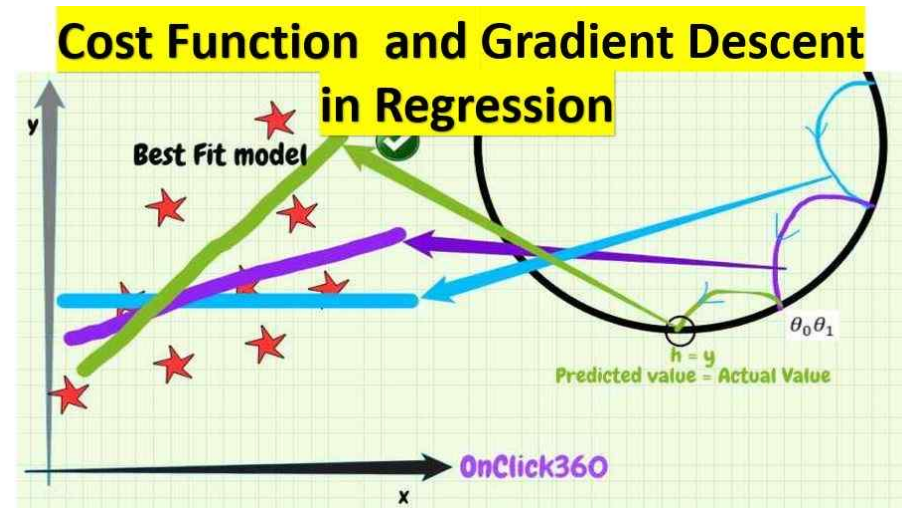
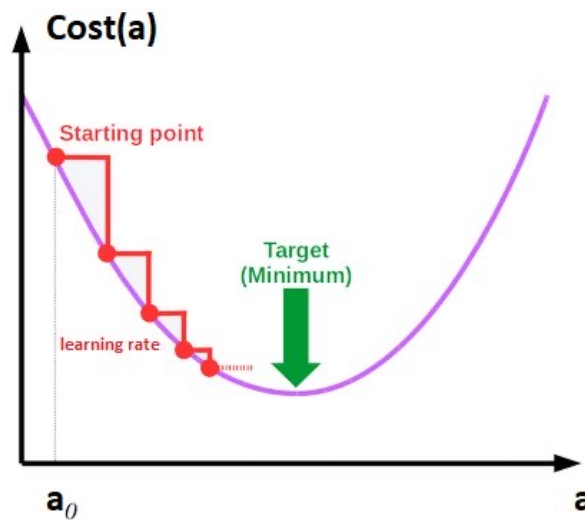
Why do we call it Machine Learning?

- ⦿ The number of data points can be very large!
- ⦿ The number of possible regression lines can be very large!
- ⦿ Trying in a brute force approach can take years, even with state-of-the-art supercomputers!
- ⦿ Exact methods also need very complex & long operations to be conducted!
- ⦿ It is called AI/ML because we "tell" the machine to efficiently find the best solution
- ⦿ "tell"=implement "intelligent" algorithms!



e.g. Gradient Descent method

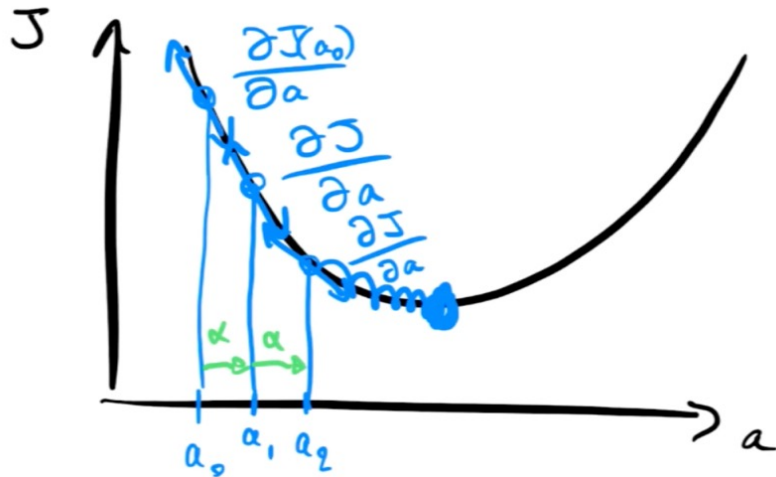
- ⦿ Gradient Descent is a method of updating a and b to reduce the error (Cost Function)
- ⦿ The idea is to start with arbitrary values for a and b then change these values iteratively to reduce the cost
- ⦿ Gradient Descent helps on how to change the values



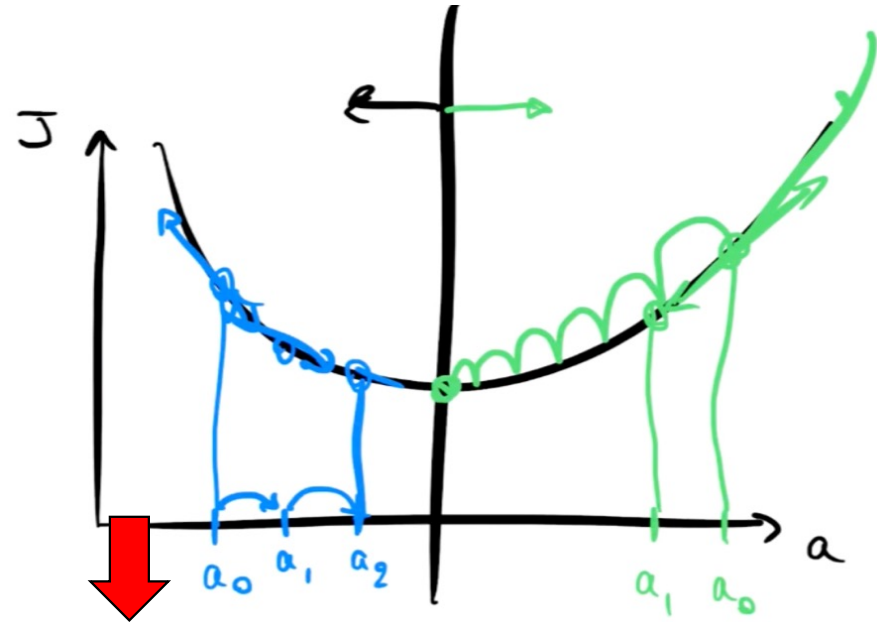
<http://aishelf.org/gradient-descent/>

Why is it more "intelligent"?

$$J(a, b) = \frac{1}{2m} \sum_{i=1}^m (f(x^{(i)}) - y^{(i)})^2$$



$$a_{i+1} = a_i - \alpha \frac{\partial J(a_i)}{\partial a}$$



$$\frac{\partial J(a_i)}{\partial a} < 0, \text{ so } -\alpha \frac{\partial J(a_i)}{\partial a} > 0$$

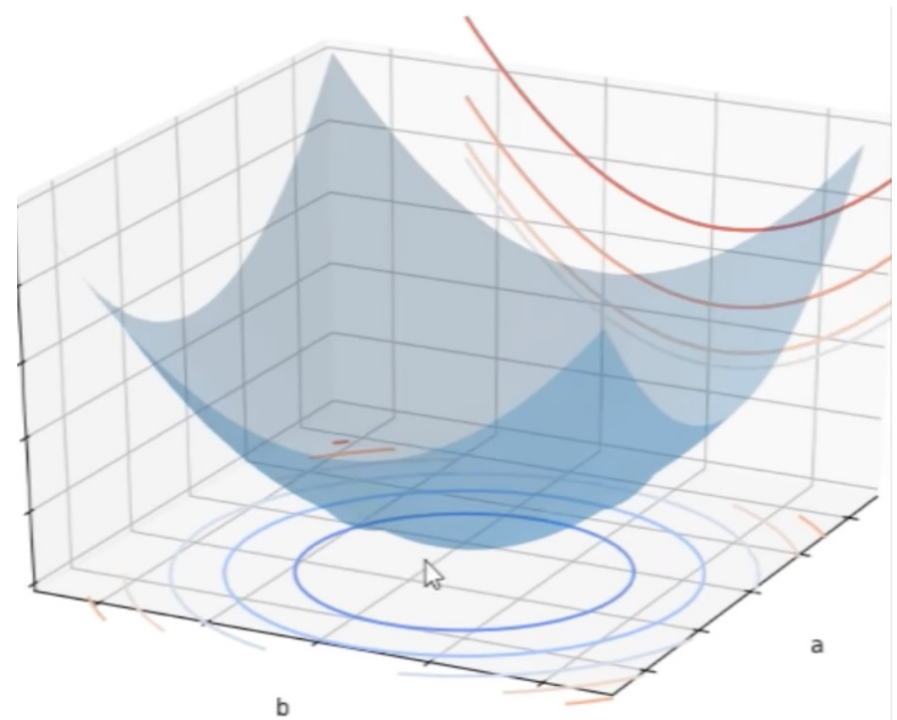
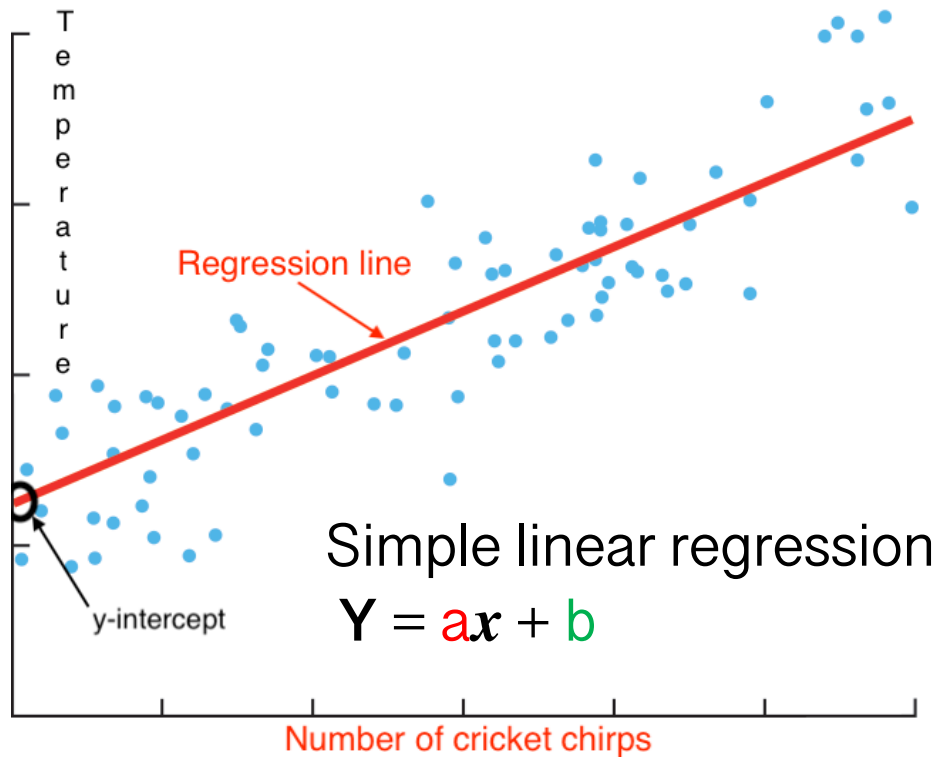
$$\text{So } a_1 > a_0$$

The algorithm converges automatically!

Why do we need that "intelligence"?

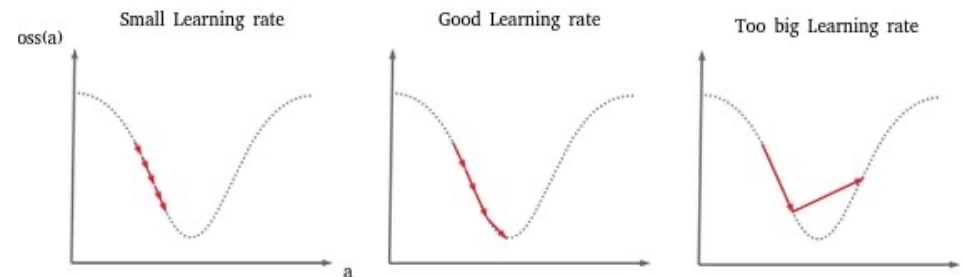
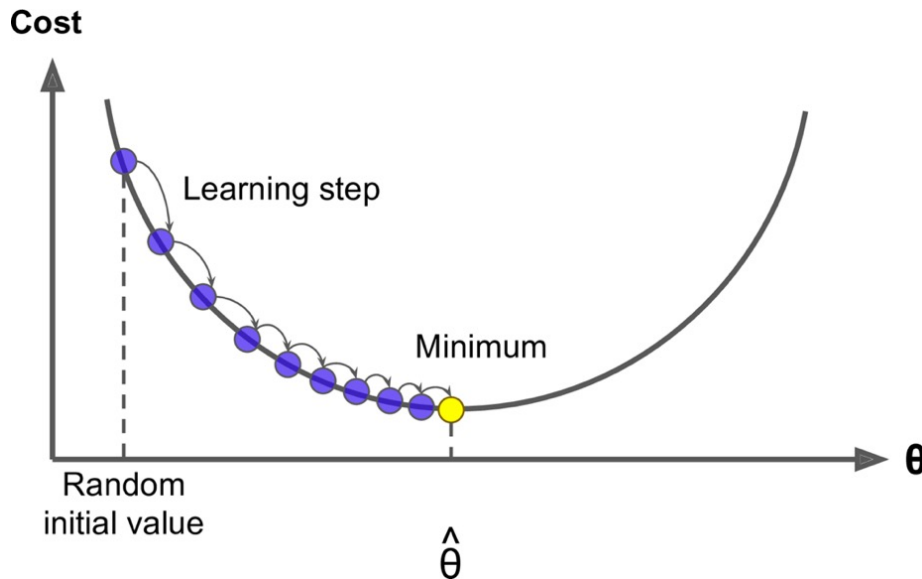
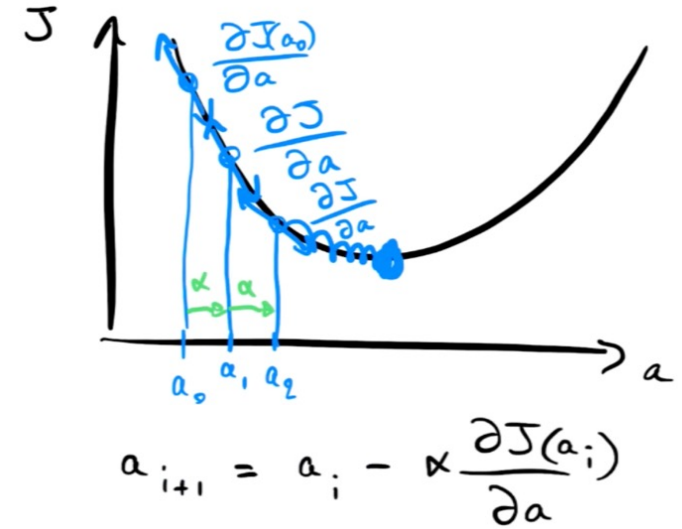
- With a simple line, we need to search for both a and b
- That already gives a large surface!
- So the search space can be huge!

Pr. Congduc Pham
<http://www.univ-pau.fr/~cpham>



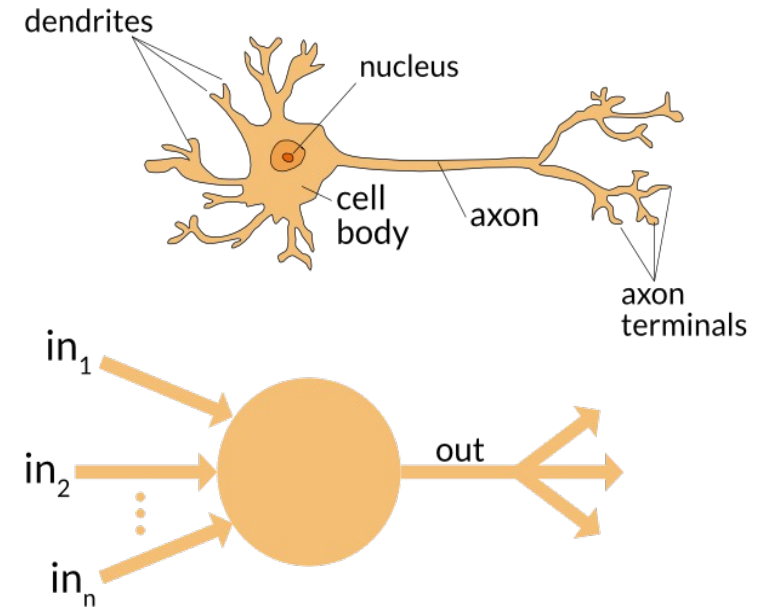
Learning Rate α

- ⦿ A smaller *learning rate* could get closer to the *minima* but takes more time to reach the *minima*
- ⦿ A larger *learning rate* converges sooner but there is a chance that you could overshoot the *minima*.



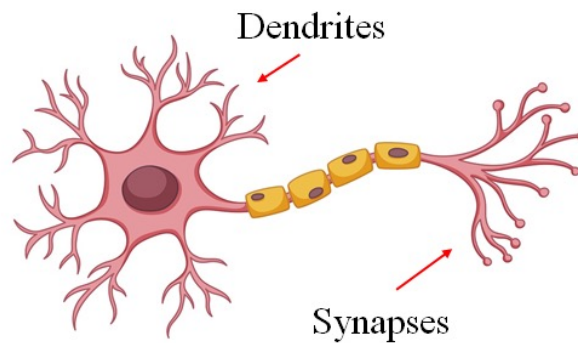
More like humans?

- ⦿ In early AI ages, researchers wanted to mimic the human brain seen as a network of neurons
- ⦿ **Perceptron**: Mathematical representation of a biological neuron
- ⦿ First implementation by Frank Rosenblatt in the 1950s
- ⦿ Rosenblatt's perceptron is activated when there is sufficient stimuli or input
- ⦿ Neurons have been found to perform a similar process, in which experience strengthens or weakens dendrites' connections

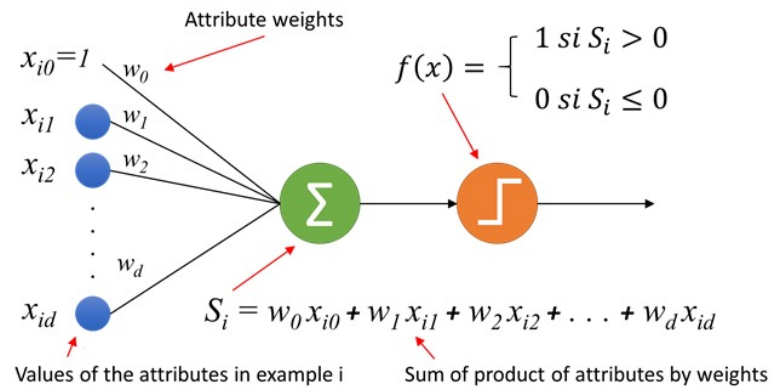


How does a Perceptron work?

- Perceptron receives the value of the attributes of an input, just as dendrites do in a neuron.
- Each attribute has a **weight** that measures its *contribution* to the final result, which is the sum of the multiplications of inputs of each attribute by its corresponding weight.
- If the sum is greater than zero Perceptron returns a value of 1, otherwise it yields 0.



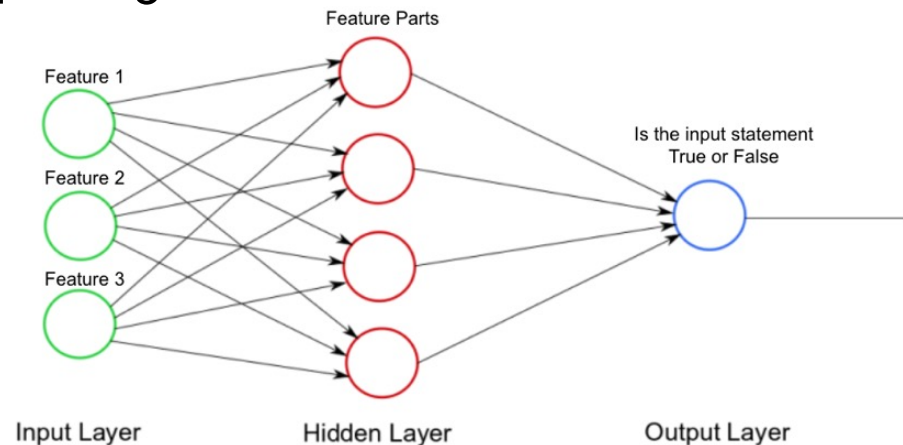
NEURON



PERCEPTRON

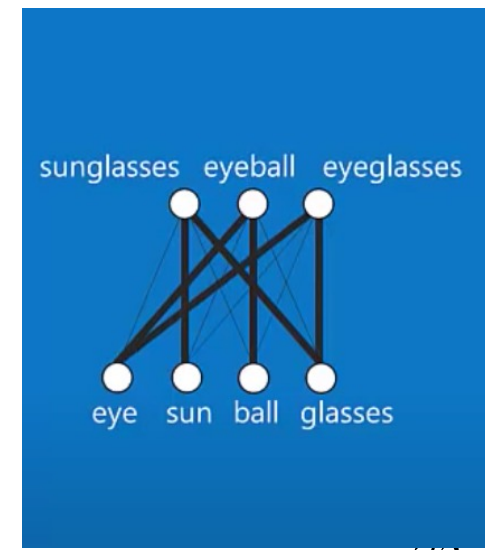
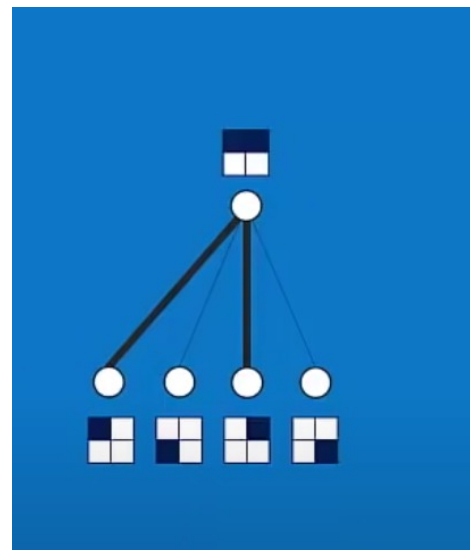
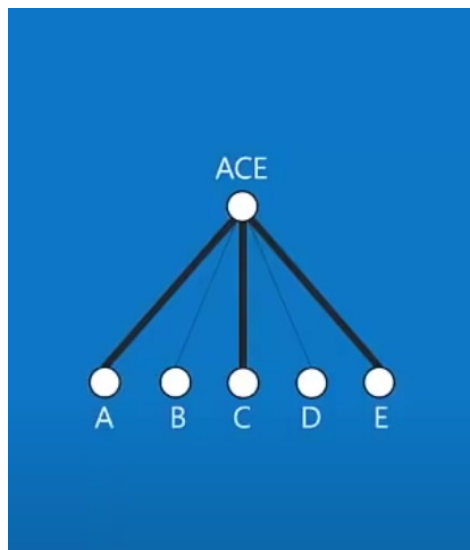
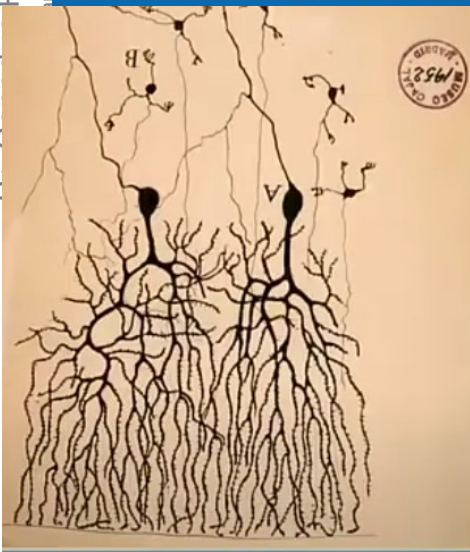
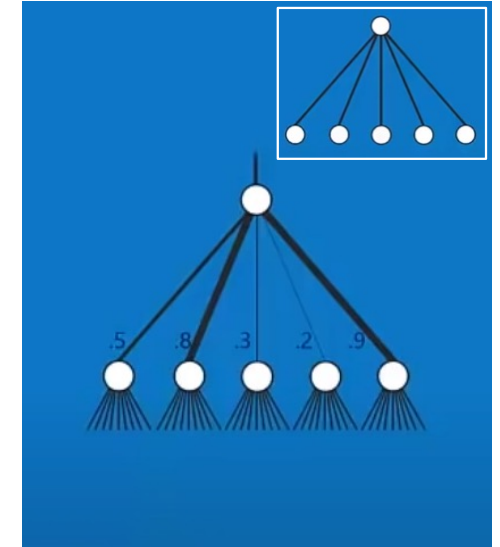
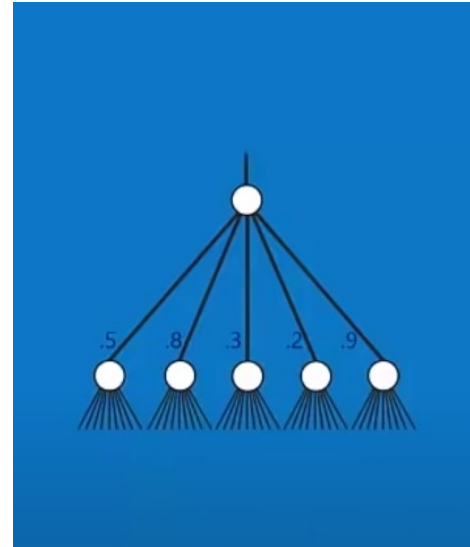
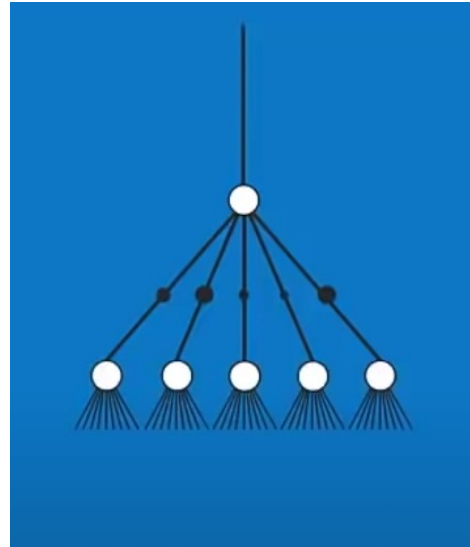
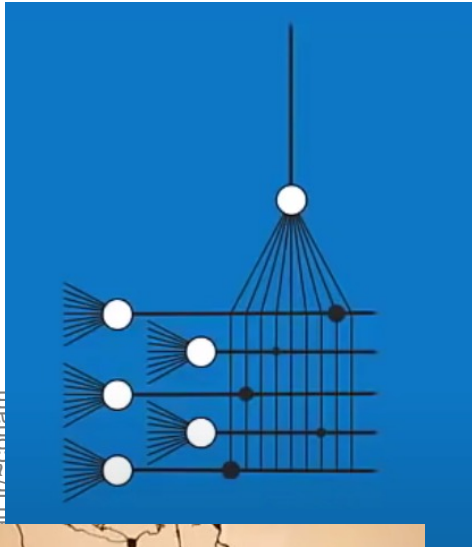
Neural Networks

- Neurons by themselves are kind of useless, in large groups, they work together to create some serious magic!
- Neural Networks are no more than a **stacking** of multiple *perceptrons* in layers to produce an output.
- Input into one layer that creates an output which in turn becomes the input for the next layer, and so on. This happens until the final output signal.



Neural Network in a nutshell

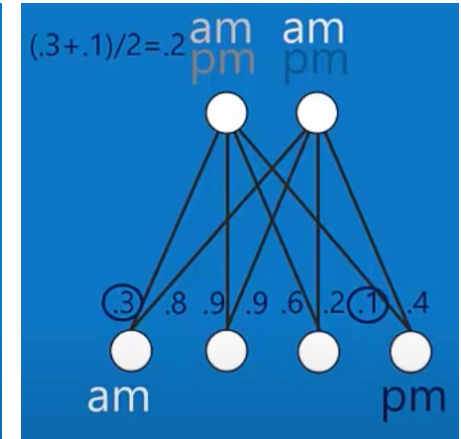
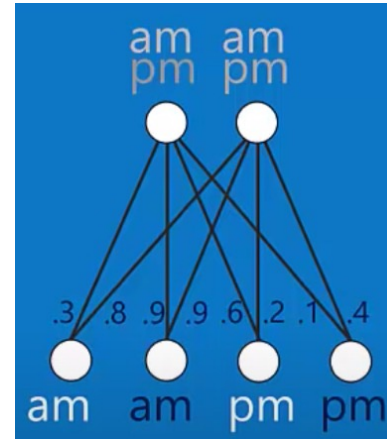
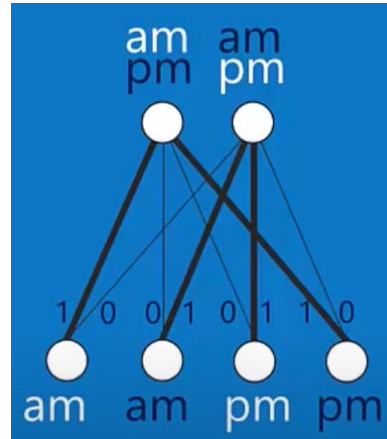
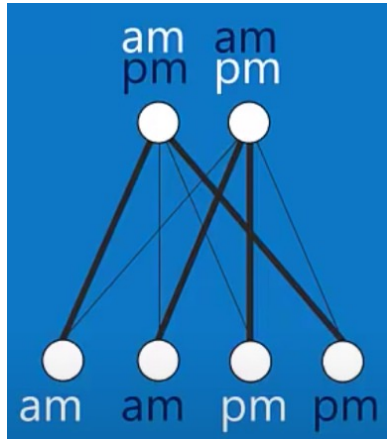
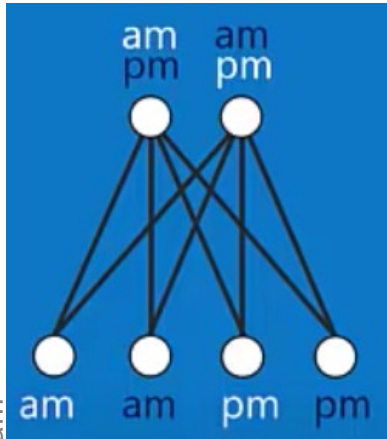
Pham
u.fr/~cpham



Pictures from <https://www.youtube.com/watch?v=Q9Z20HCPnww> (Brandon Rohrer)

Shawarma guy example (1)

Pictures from <https://www.youtube.com/watch?v=Q9Z20HCPnww> (Brandon Rohrer)

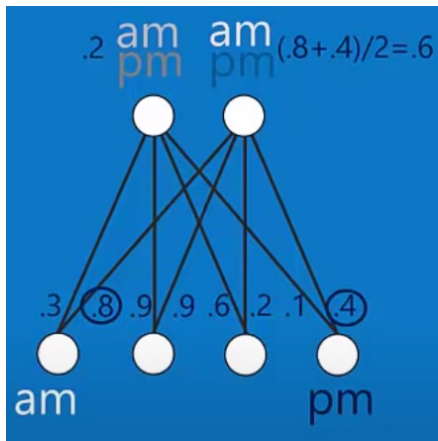


Pr. Congduc Pham
http://www.unipr.edu.vn/~pcpam/

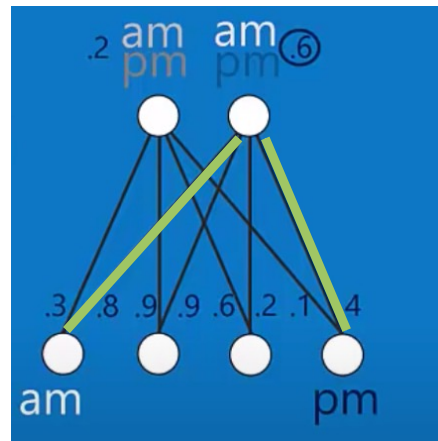
Sometimes he works mornings,
sometimes evenings

Start with random weights

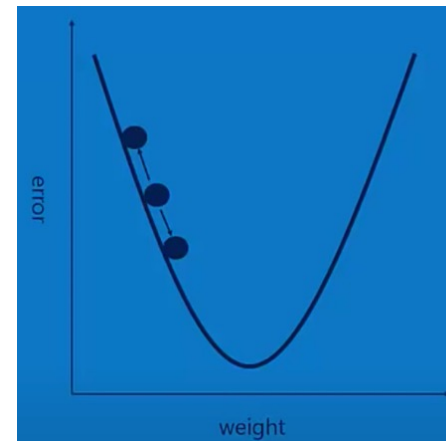
Observed on a given day



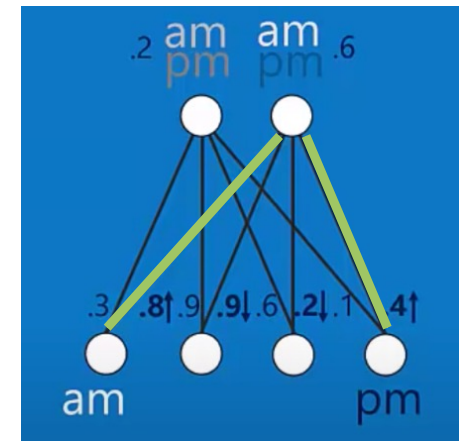
Error = 1 - .2 = .8



Error = 1 - .6 = .4

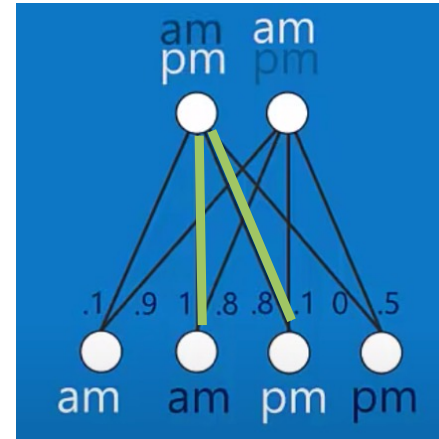
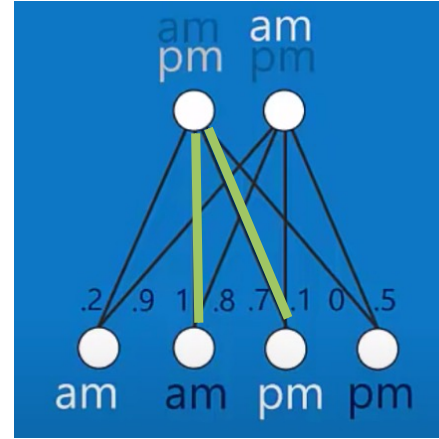
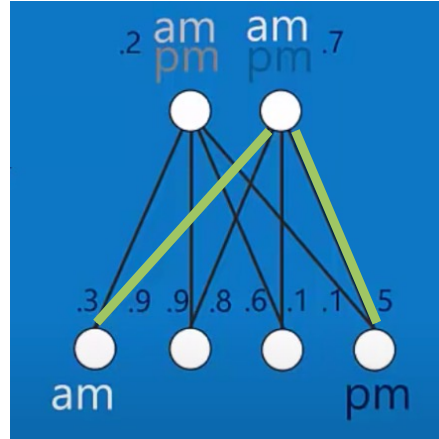
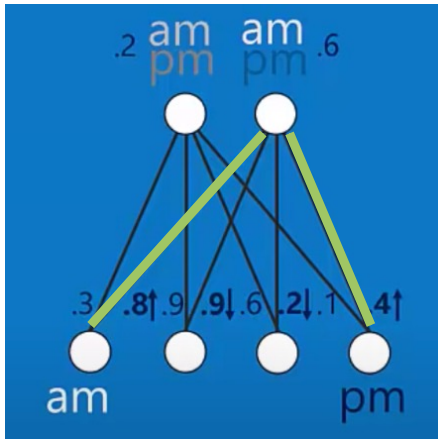


Here comes again our Gradient Descent methods!



Shawarma guy example (2)

Pictures from <https://www.youtube.com/watch?v=Q9Z20HCPnww> (Brandon Rohrer)

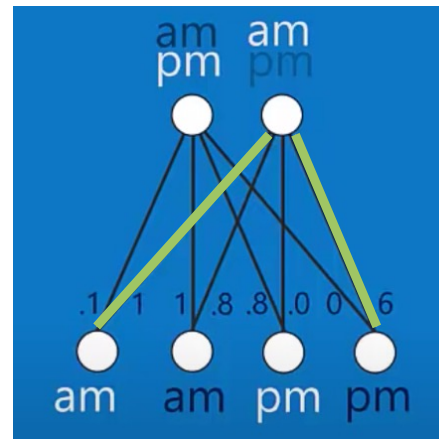


am pm Increase active links, decrease inactive links

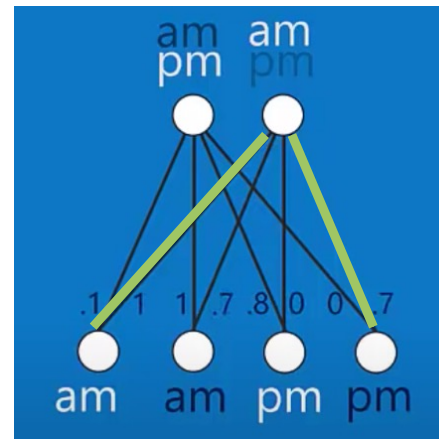
Error = $1 - .7 = .3$

am pm The next day

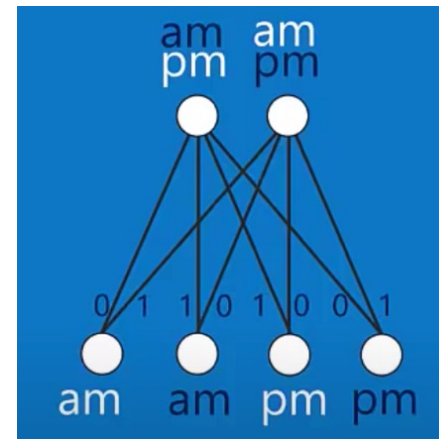
am pm The next day



am pm The next day



am pm The next day.....



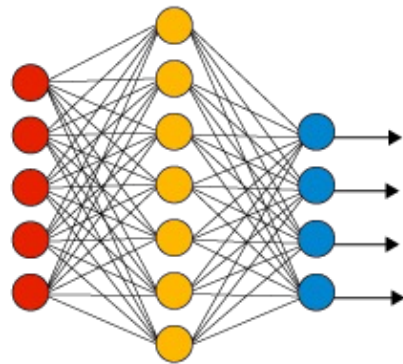
Eventually, weights don't change anymore!

At this point you have a trained model!

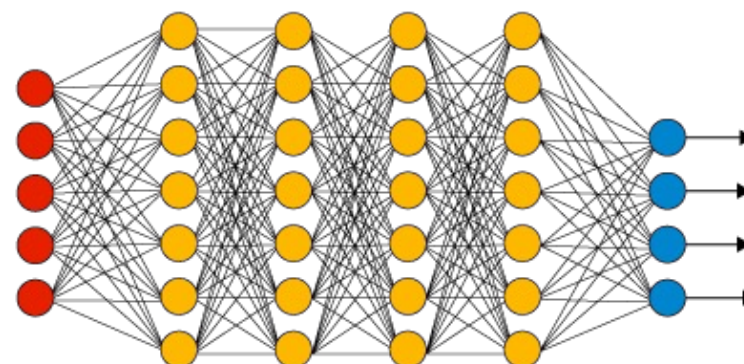
Deep Neural Networks

- ⦿ In the 1980s, most Artificial Neural Networks were single-layered due to the cost of computation and availability of data.
- ⦿ Nowadays is possible to afford more hidden layers, hence the moniker “Deep Neural Networks”.
- ⦿ Regained popularity since ~2006 and rebranded as Deep Learning (DL)

Simple Neural Network



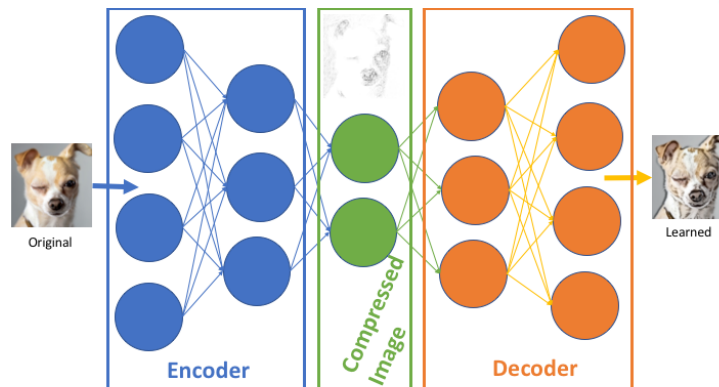
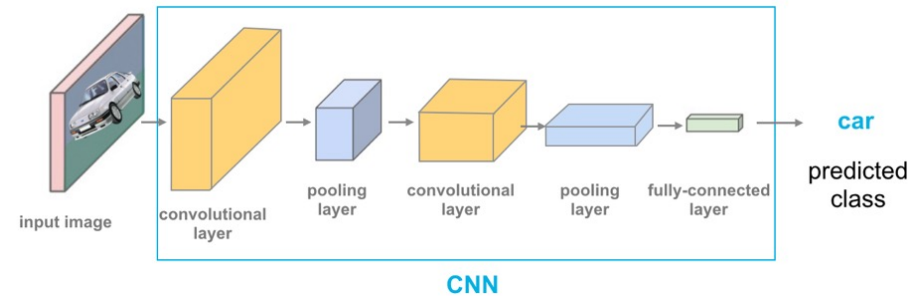
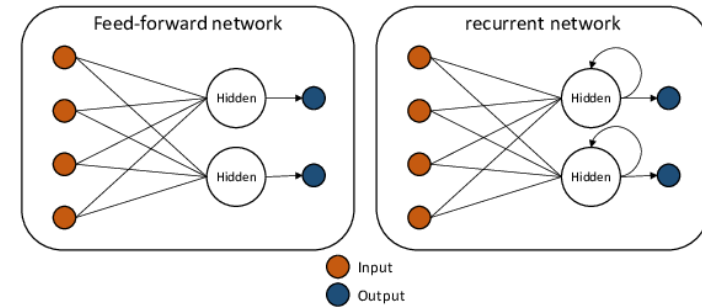
Deep Learning Neural Network



● Input Layer ● Hidden Layer ● Output Layer

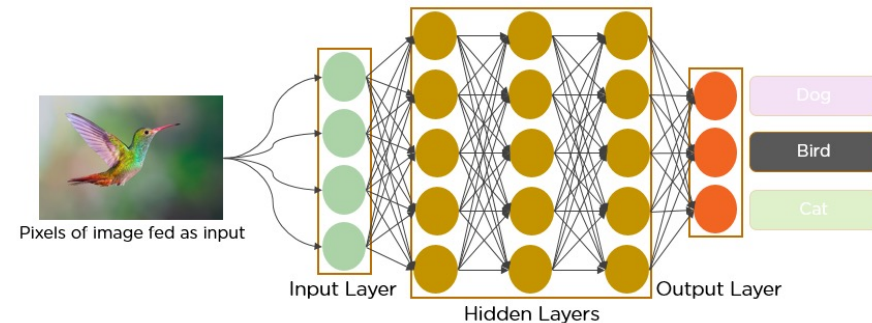
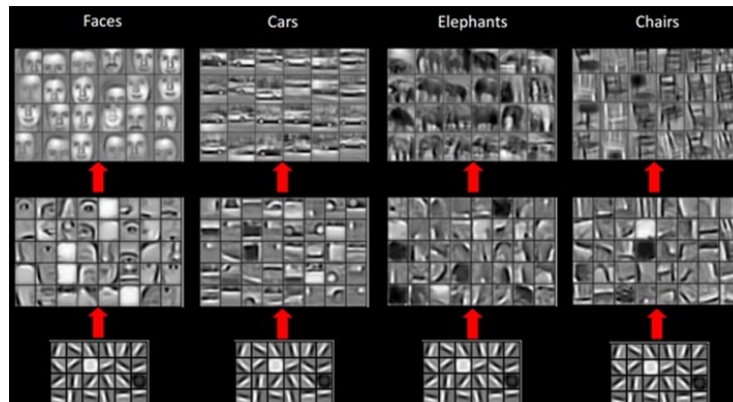
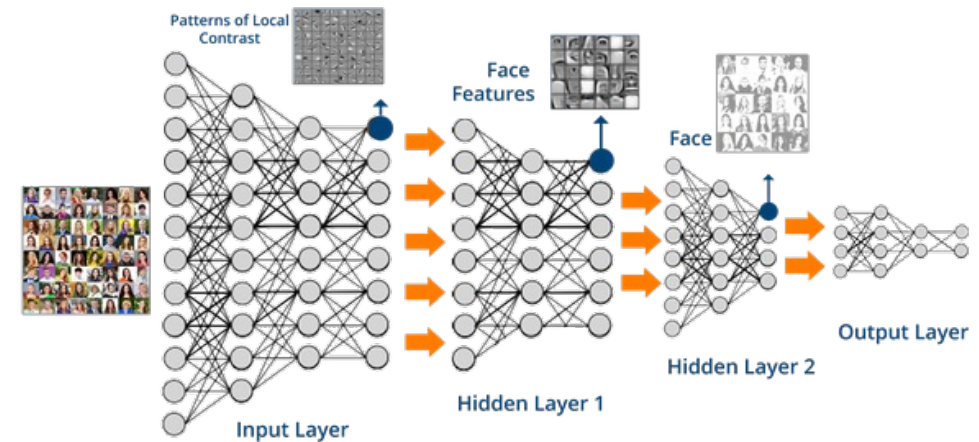
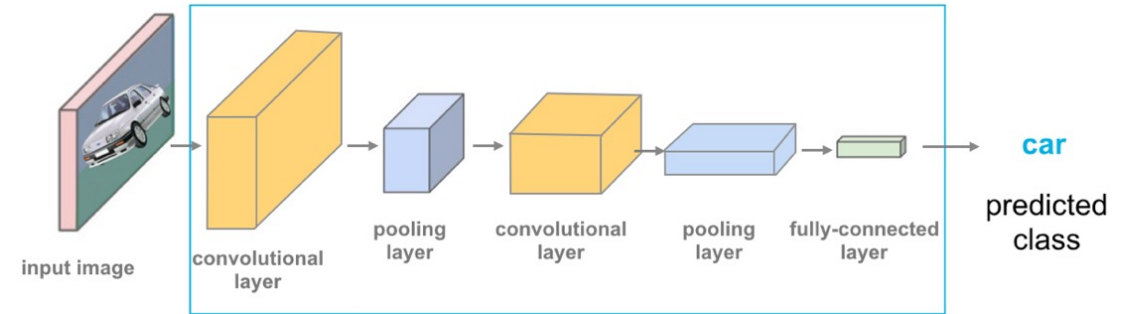
Types of Deep Neural Networks

- Feedforward Neural Networks (**FFNs, ANNs** or **NNs**)
- Recurrent Neural Networks (**RNNs**)
- Convolutional Neural Networks (**CNNs**)
- Autoencoder Neural Networks (**AEs**)

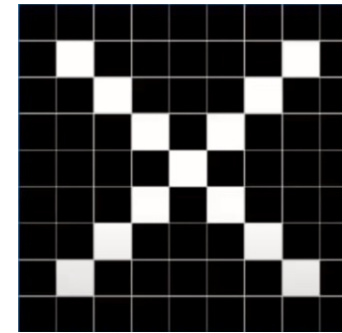
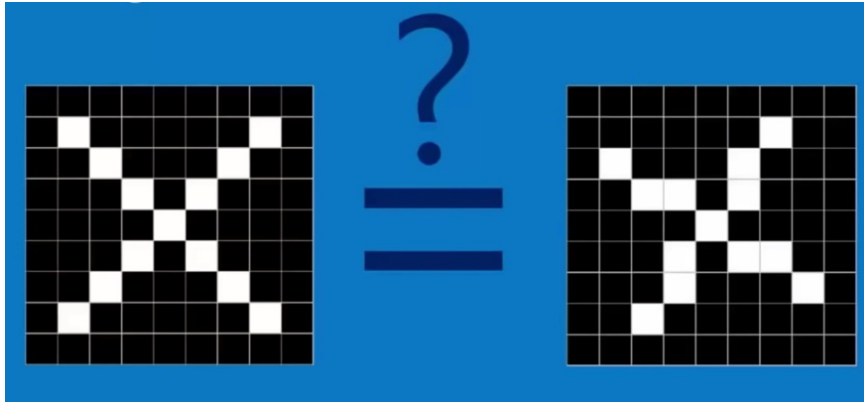


Convolutional Neural Networks

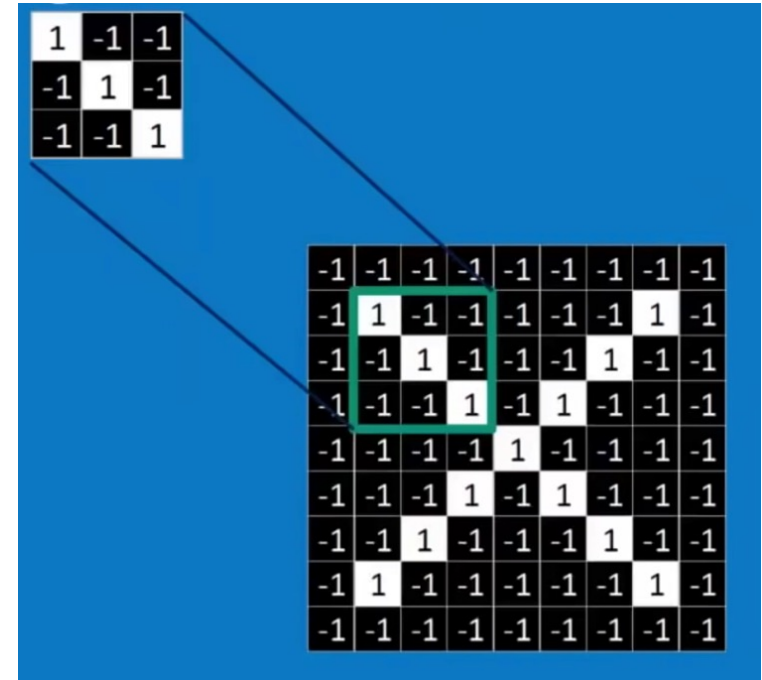
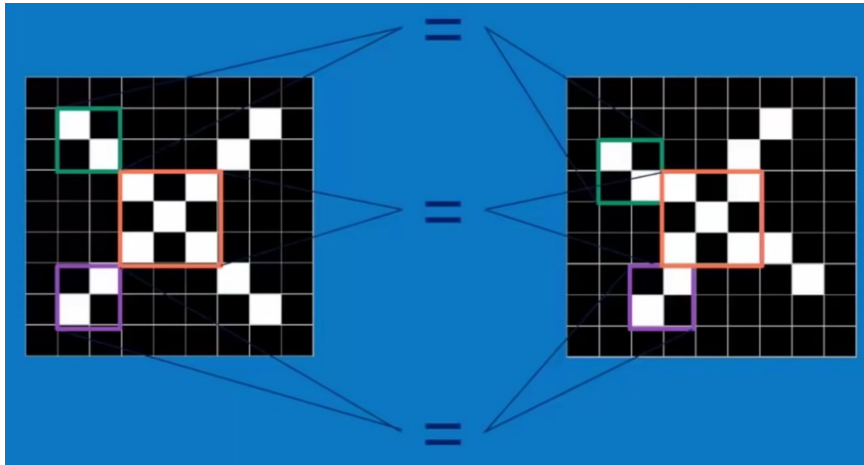
- Contain five types of layers
- Each layer has a specific purpose, like summarizing, connecting or activating
- CNN are good at image classification and object detection



What is a convolution?



-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

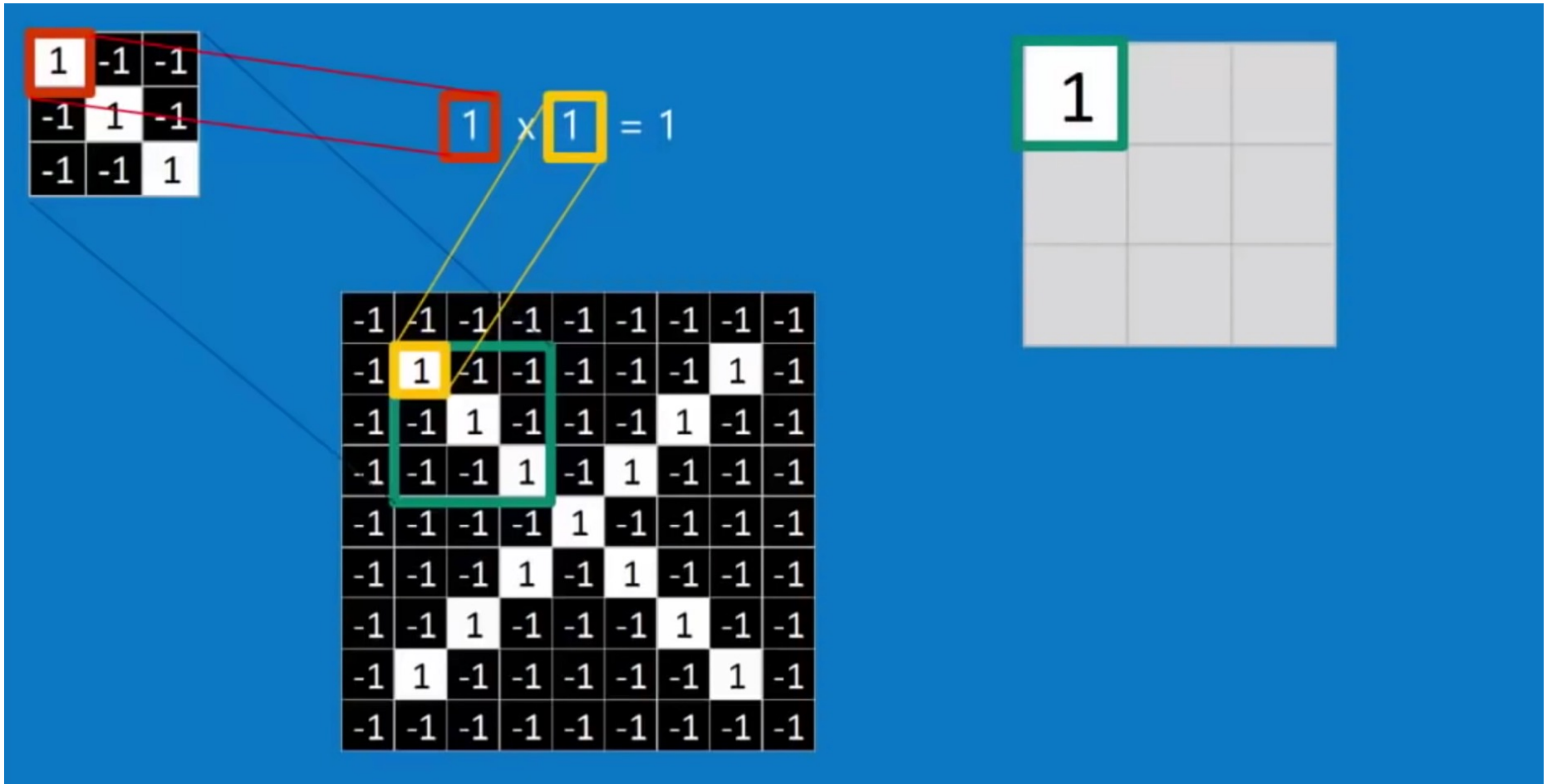


1	-1	-1
-1	1	-1
-1	-1	1

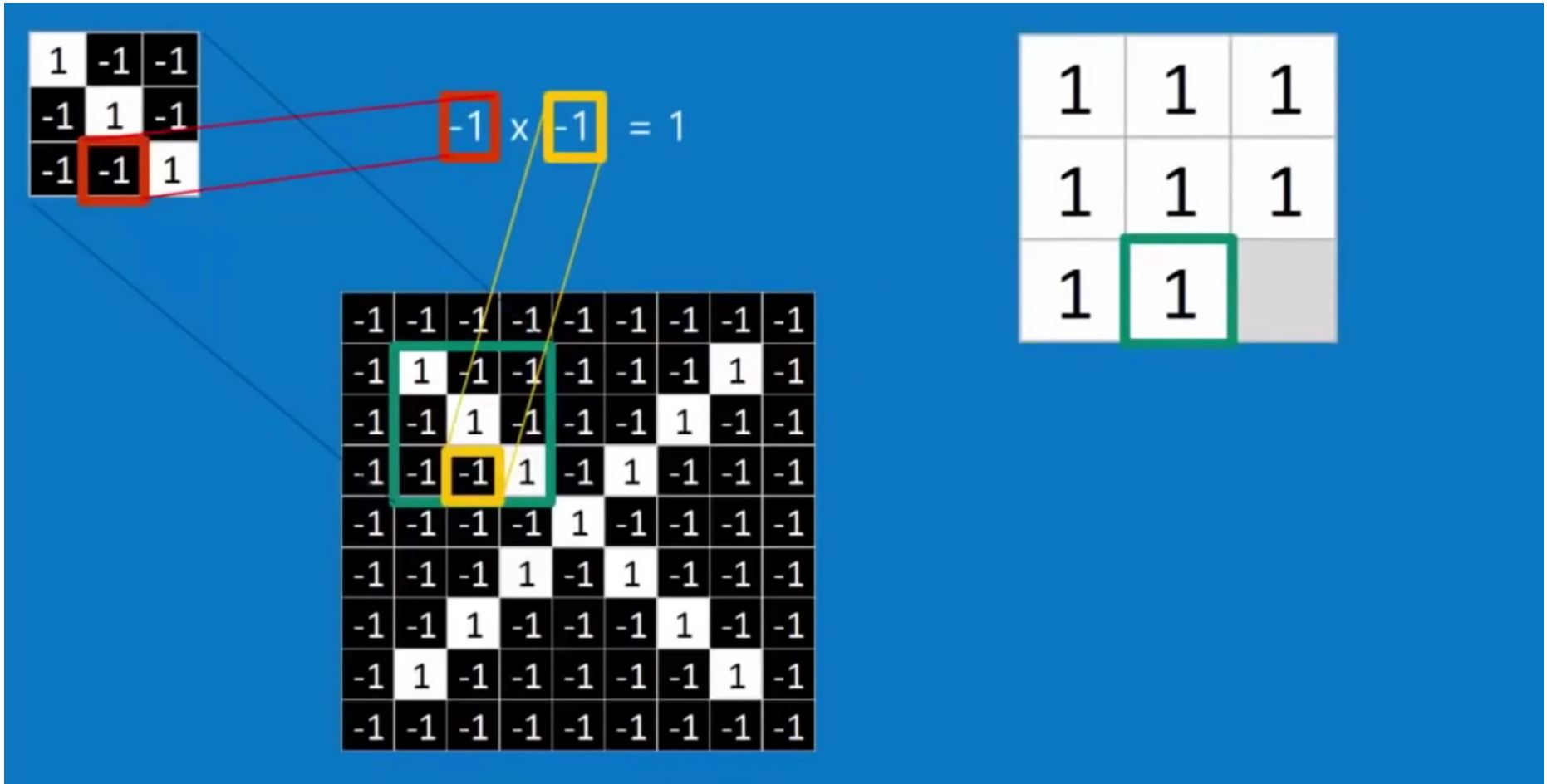
1	-1	1
-1	1	-1
1	-1	1

-1	-1	1
-1	1	-1
1	-1	-1

Running the convolution (1)



Running the convolution (2)



Running the convolution (3)

The diagram illustrates the final step of a convolution operation. It shows a 3x3 kernel, a 3x3 kernel of ones, a 9-term sum, and a 9x9 grid with a 3x3 region highlighted.

Kernel 1 (3x3):

1	-1	-1
-1	1	-1
-1	-1	1

Kernel 2 (3x3):

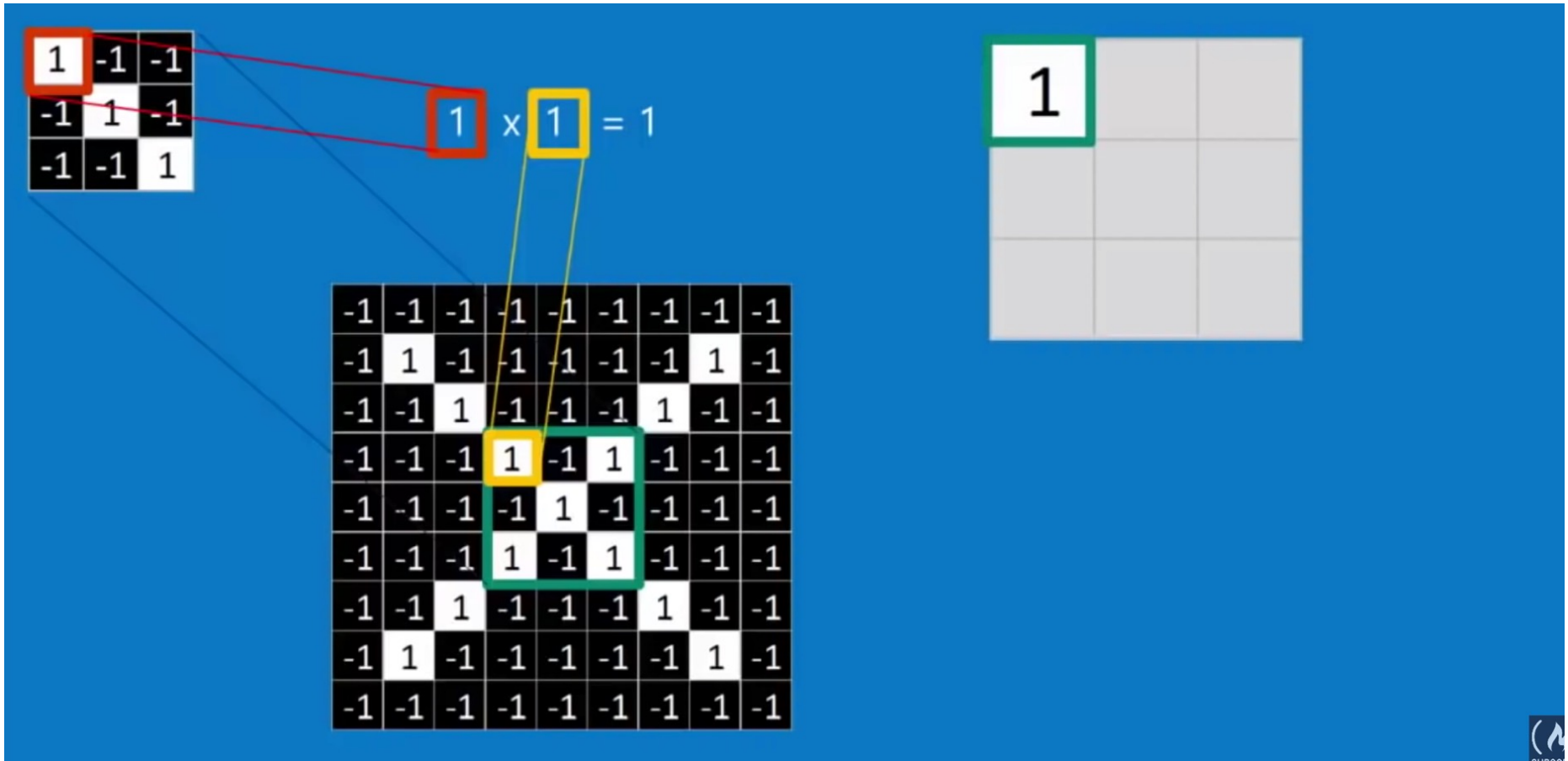
1	1	1
1	1	1
1	1	1

Sum:

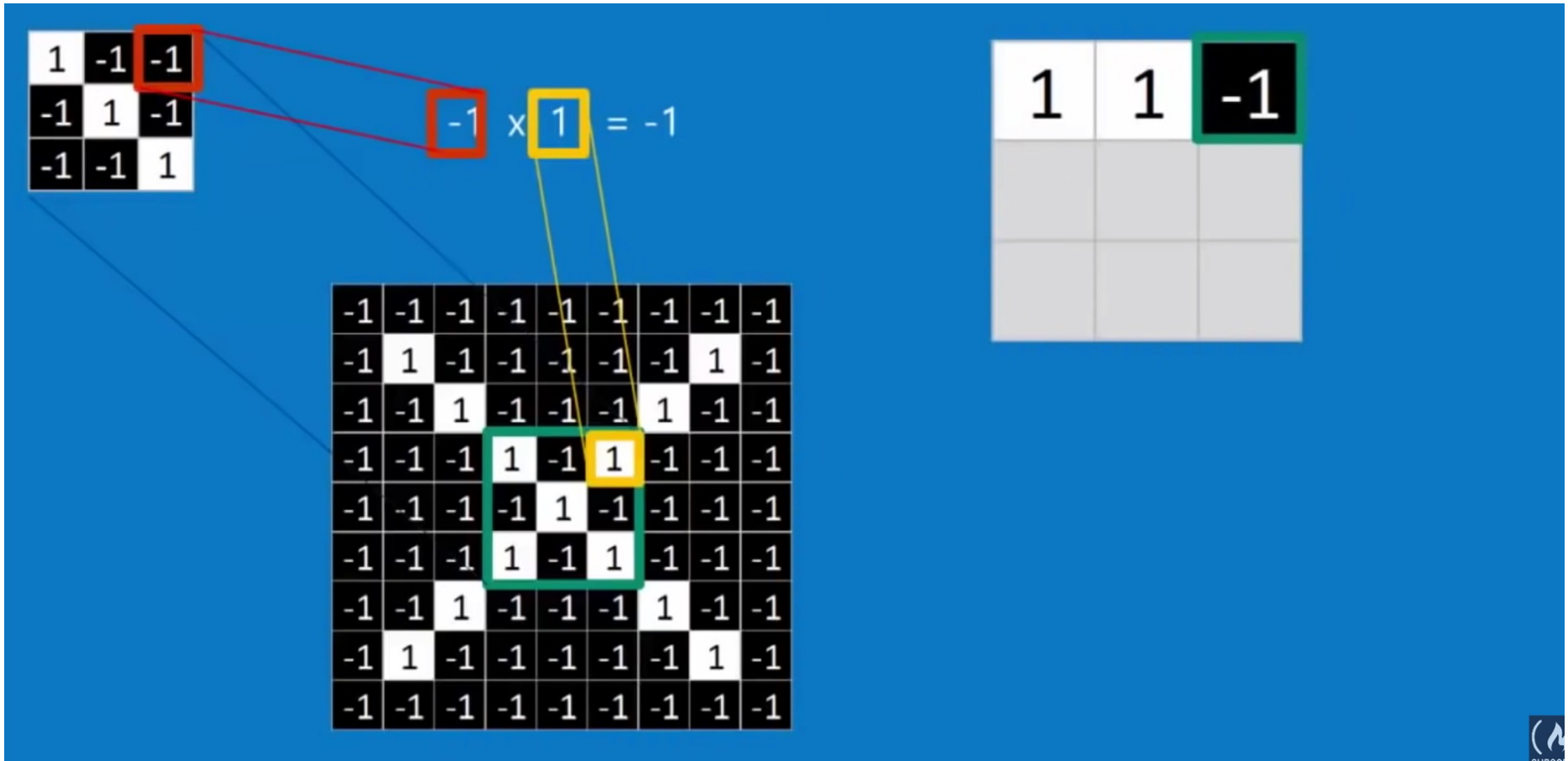
$$\frac{1+1+1+1+1+1+1+1+1}{9} = 1$$

Grid: A 9x9 grid where a 3x3 region is highlighted in green. The value 1 is shown in the center of this region.

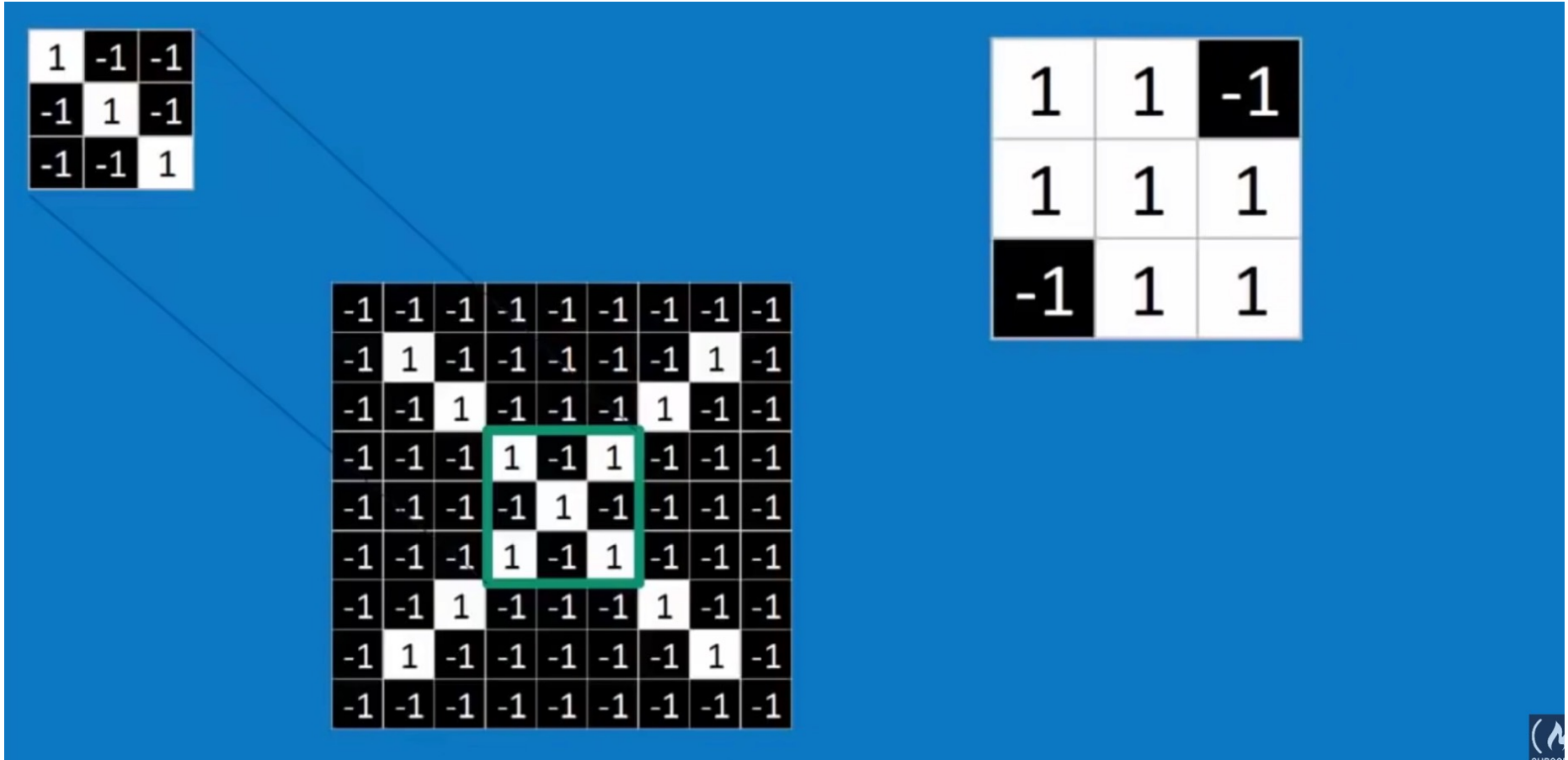
Running the convolution (4)



Running the convolution (5)



Running the convolution (6)



Running the convolution (7)

The diagram illustrates a 3x3 convolution operation on a 9x9 grid. A 3x3 kernel is applied to a 3x3 region of the input grid to produce a single output value of 0.55. The calculation is shown as $\frac{1+1-1+1+1+1-1+1+1}{9} = 0.55$.

Kernel 1 (top left):

1	-1	-1
-1	1	-1
-1	-1	1

Kernel 2 (top right):

1	1	-1
1	1	1
-1	1	1

Input Grid (middle):

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

Output Grid (right):

Calculation:

$$\frac{1+1-1+1+1+1-1+1+1}{9} = 0.55$$

Running the convolution (8)

Pr. Congduc Pham
<http://www.univ-pau.fr/~cpham>

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1



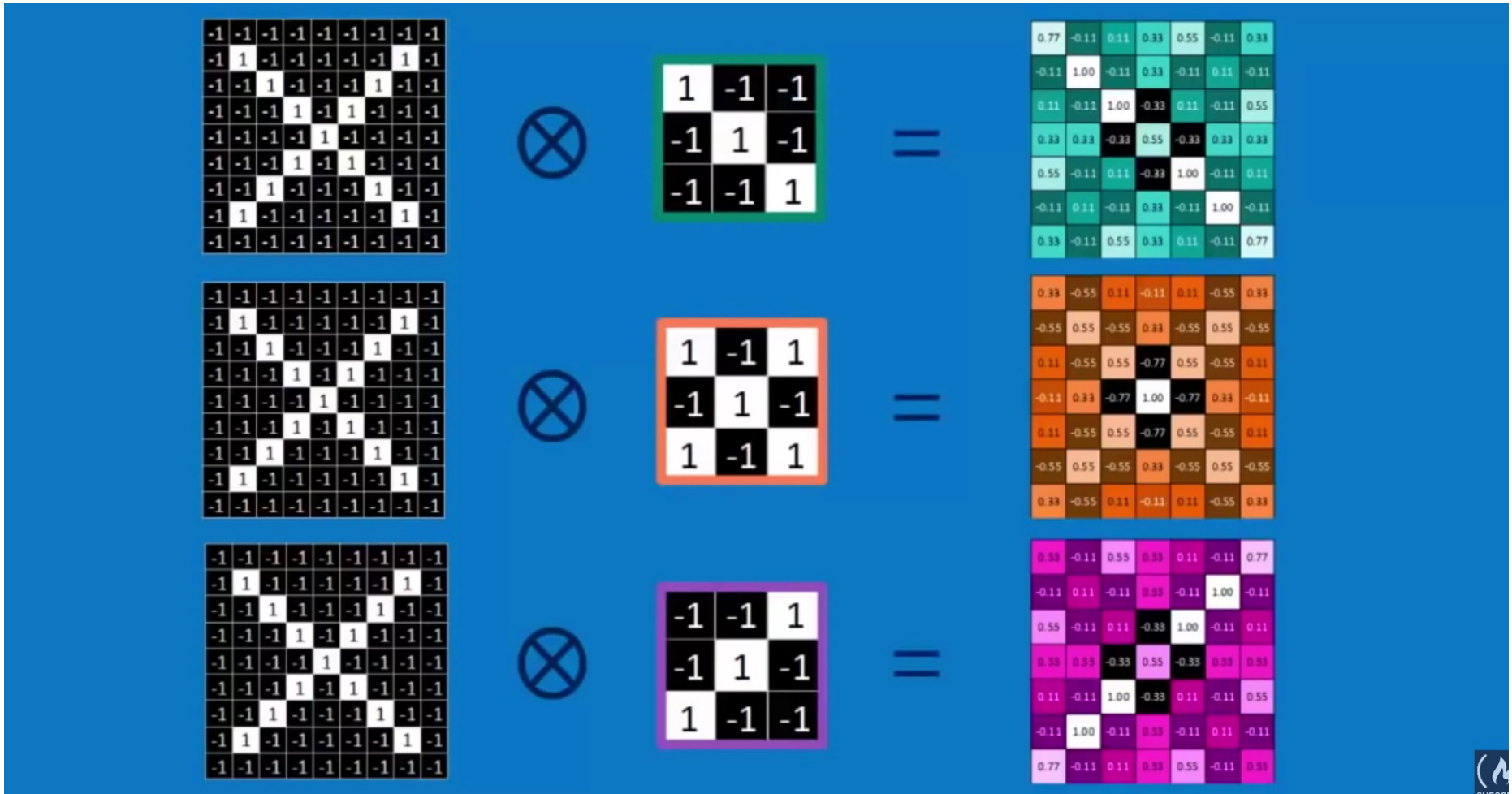
1	-1	-1
-1	1	-1
-1	-1	1



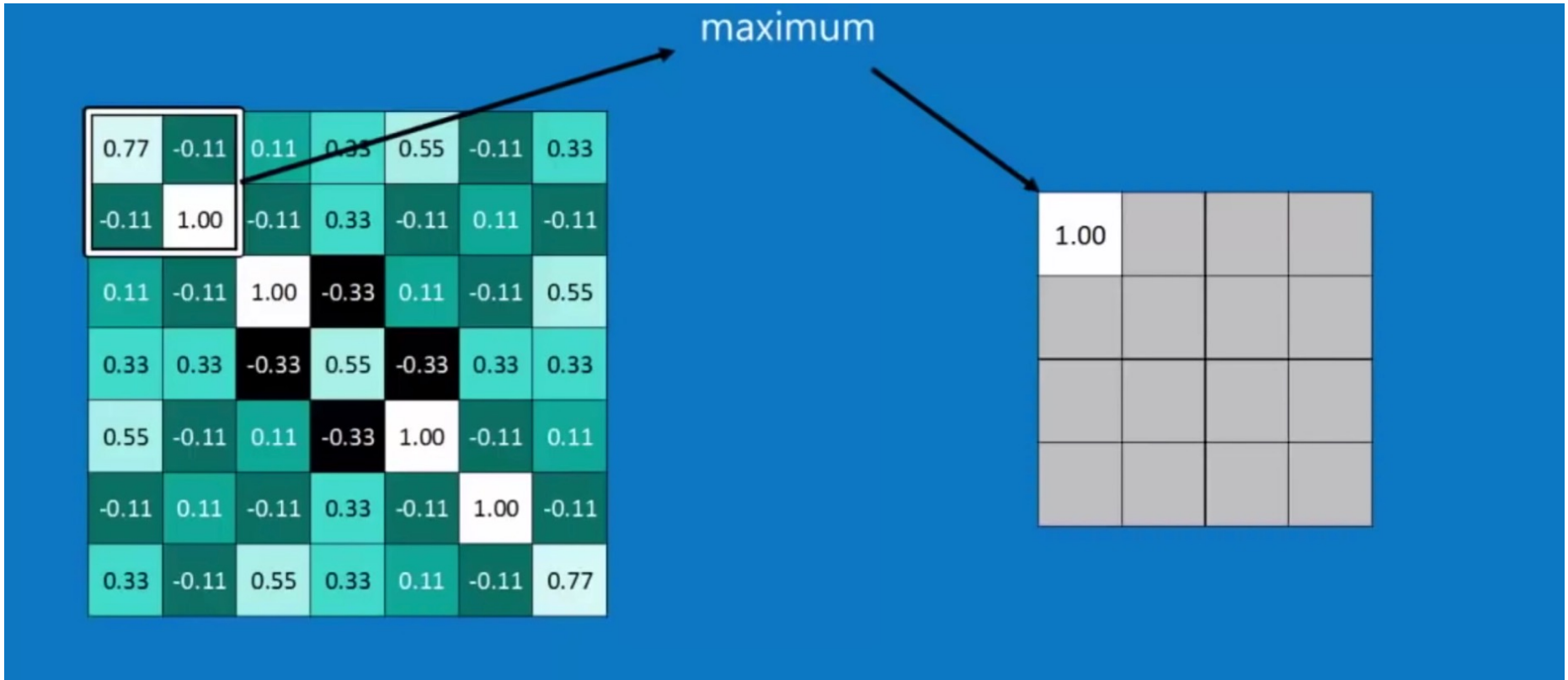
0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77



1 image = stack of filtered images



Next, Pooling = reduce the size



At the end of pooling

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77

max pooling

1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

Do the same for stack of images

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77

0.33	-0.55	0.11	-0.11	0.11	-0.55	0.33
-0.55	0.55	-0.55	0.33	-0.55	0.55	-0.55
0.11	-0.55	0.55	-0.77	0.55	-0.55	0.11
-0.11	0.33	-0.77	1.00	-0.77	0.33	-0.11
0.11	-0.55	0.55	-0.77	0.55	-0.55	0.11
-0.55	0.55	-0.55	0.33	-0.55	0.55	-0.55
0.33	-0.55	0.11	-0.11	0.11	-0.55	0.33

0.33	-0.11	0.55	0.33	0.11	-0.11	0.77
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.77	-0.11	0.11	0.33	0.55	-0.11	0.33



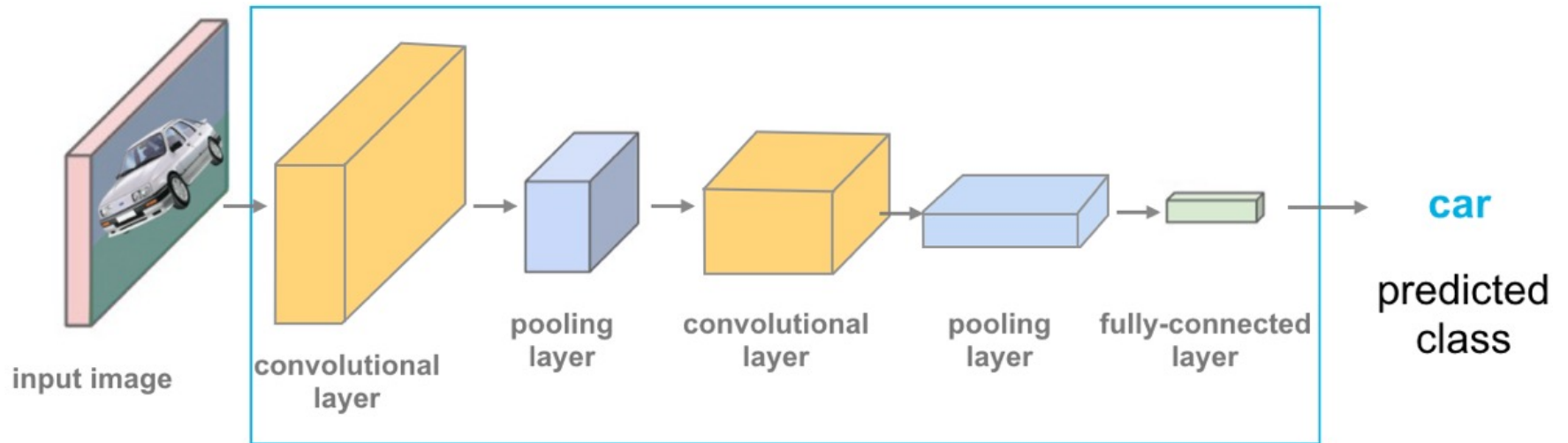
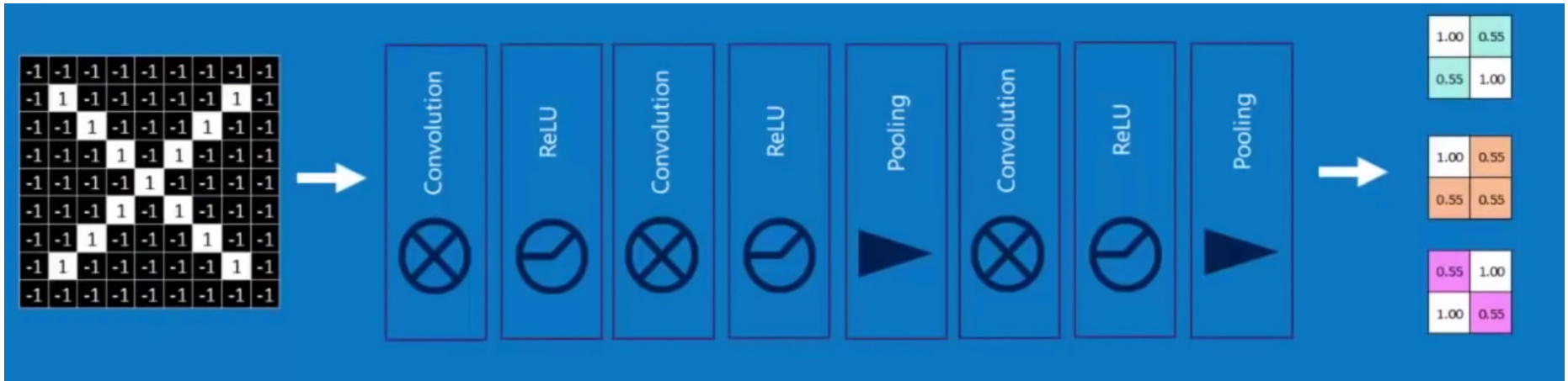
1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

0.55	0.33	0.55	0.33
0.33	1.00	0.55	0.11
0.55	0.55	0.55	0.11
0.33	0.11	0.11	0.33

0.33	0.55	1.00	0.77
0.55	0.55	1.00	0.33
1.00	1.00	0.11	0.55
0.77	0.33	0.55	0.33

Layers can be repeated!

Watch the video on YouTube <https://www.youtube.com/watch?v=FmpDlaiMleA> (Brandon Rohrer)



Want to try DeepLearning?

🕒 <https://playground.tensorflow.org/>

Tinker With a **Neural Network** Right Here in Your Browser.
Don't Worry, You Can't Break It. We Promise.

🔄 ▶️ Epoch: 000,000 Learning rate: 0.03 Activation: Tanh Regularization: None Regularization rate: 0 Problem type: Classification

DATA

Which dataset do you want to use?

Ratio of training to test data: 50%

Noise: 0

Batch size: 10

REGENERATE

FEATURES

Which properties do you want to feed in?

- X_1
- X_2
- X_1^2
- X_2^2
- $X_1 X_2$
- $\sin(X_1)$
- $\sin(X_2)$

2 HIDDEN LAYERS

4 neurons 2 neurons

The outputs are mixed with varying weights, shown by the thickness of the lines.

This is the output from one neuron. Hover to see it larger.

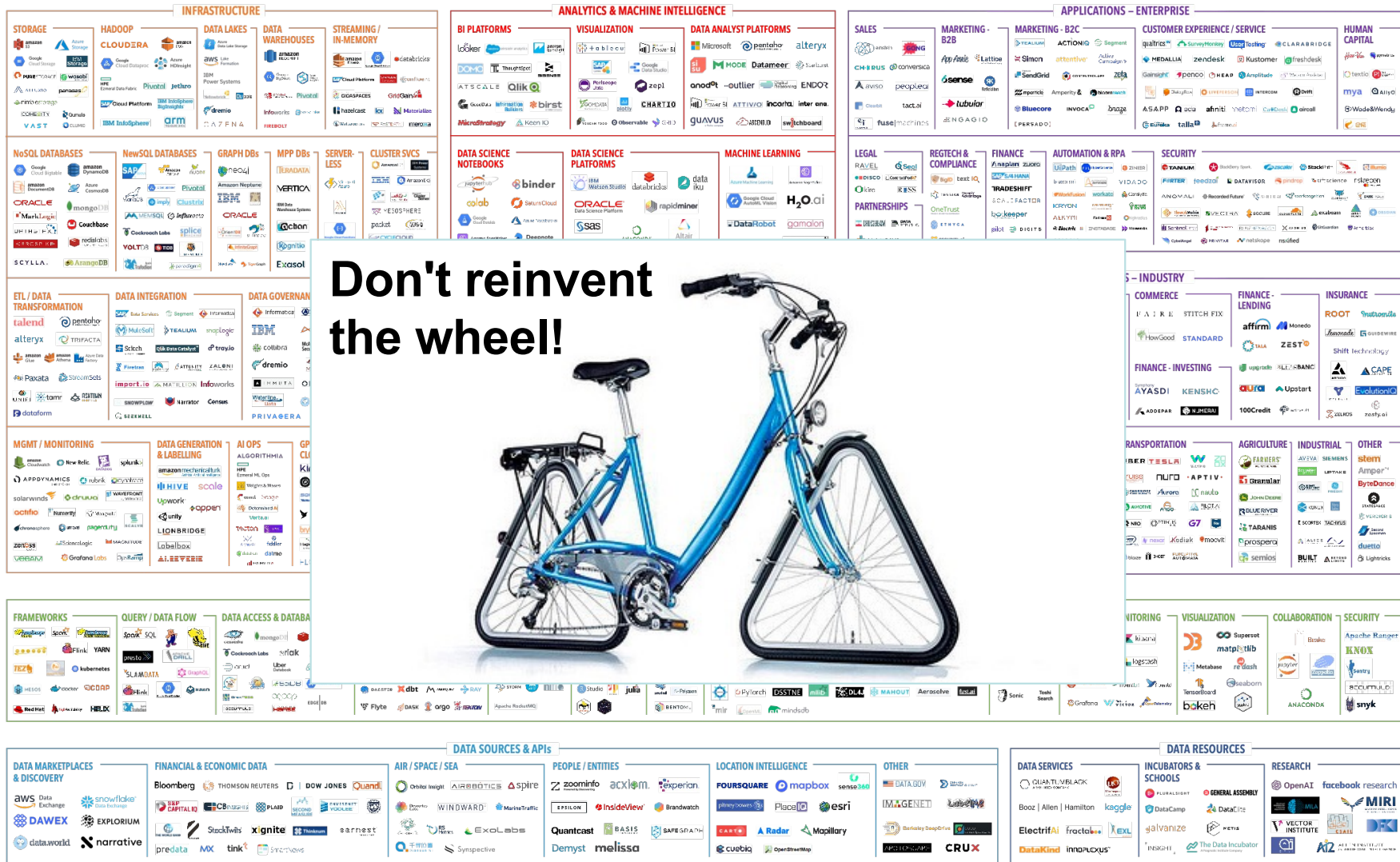
OUTPUT

Test loss 0.522
Training loss 0.508

Colors shows data, neuron and weight values.

The BigData & AI Landscape

DATA & AI LANDSCAPE 2020



Don't reinvent the wheel!

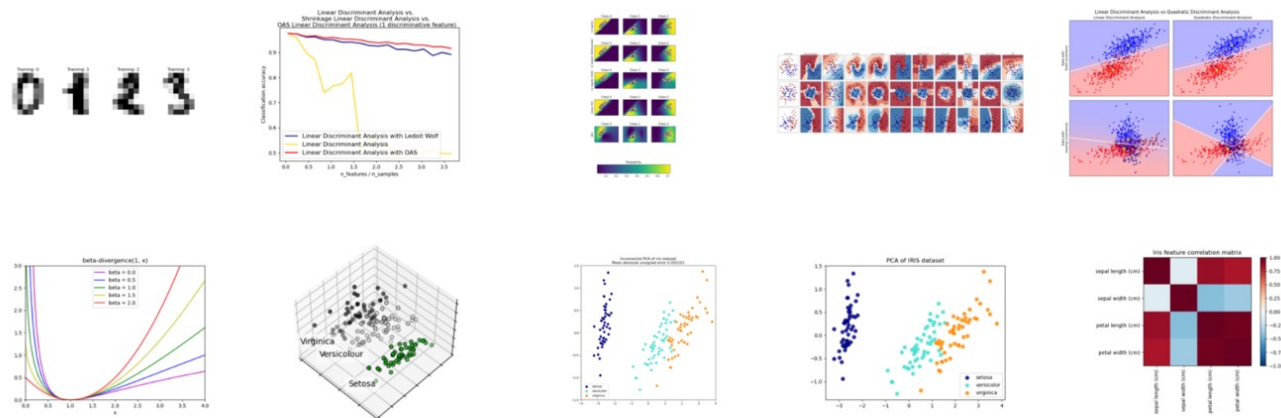
Pr. Congduc Pham
http://www.univ-pau.fr/~cpham

Package: pandas

- ⦿ Python Data Analysis Library
- ⦿ pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language
- ⦿ pandas provides fast, flexible, and expressive data structures designed to make working with “relational” or “labeled” data both easy and intuitive
- ⦿ it aims to be the fundamental high-level building block for doing practical, real-world data analysis in Python.
- ⦿ https://youtu.be/_T8LGqJtuGc

Package: Scikit-Learn

- Scikit-learn is an open source machine learning library that supports supervised and unsupervised learning
- It also provides various tools for model fitting, data preprocessing, model selection and evaluation, and many other utilities
- https://scikit-learn.org/stable/auto_examples/index.html



Package: TensorFlow

- ⦿ TensorFlow is an end-to-end open source platform for machine learning and deep learning
- ⦿ It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications
- ⦿ "TensorFlow is more of a low-level library whereas Scikit-Learn comes with off-the-shelf algorithms, e.g., algorithms for classification such as SVMs, Random Forests, Logistic Regression, ..."
[<https://stackoverflow.com/questions/61233004>]
- ⦿ <https://www.tensorflow.org/>
- ⦿ <https://www.tensorflow.org/overview>

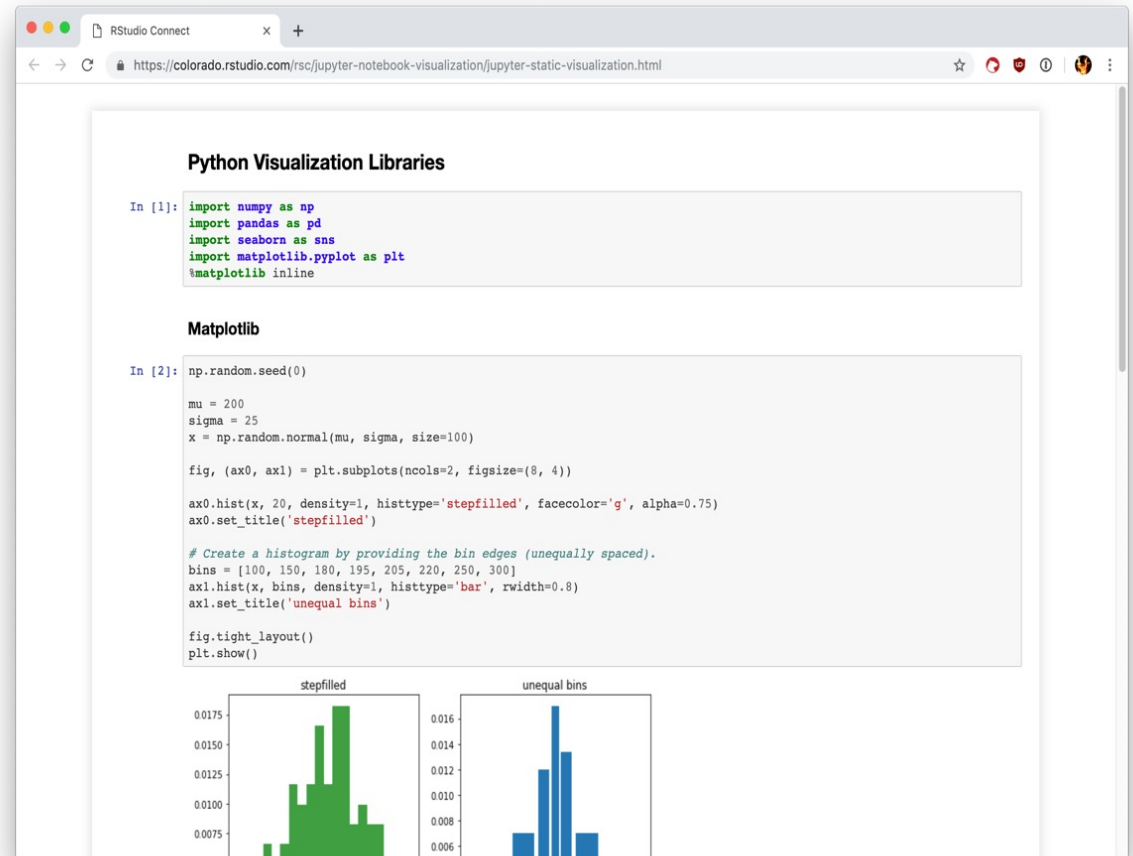
Package: XGBoost

- ⦿ eXtreme Gradient Boosting is an optimized open source implementation of the gradient boosting trees algorithm
- ⦿ Gradient Boosting is a **supervised learning algorithm** whose principle is to combine the results of a set of data and weaker models in order to provide a **better prediction** (regression)
- ⦿ XGBoost includes a large number of hyperparameters which can be modified and tuned for improvement
- ⦿ XGBoost is not part of Scikit-Learn but works perfectly with it
- ⦿ XGBoost behaves remarkably in machine learning competitions!
- ⦿ Source: "XGBoost: The super star of algorithms in ML competition". See examples from <http://aishelf.org/xgboost/>

Jupyter Notebook

Web application that allows to create and share documents that contain live code, equations, visualizations and narrative text.

- ⦿ Data cleaning and transformation
- ⦿ Numerical simulation
- ⦿ Statistical modeling
- ⦿ Data visualization
- ⦿ Machine learning, and more.



Conclusions

- ⦿ Internet-of-Things provides the unique feature to make things "talk" to us: localisation, surrounding environmental conditions, particular events, ...
- ⦿ After many years of maturing IoT technologies...
- ⦿ ... current trends is to optimize IoT for verticals
- ⦿ IoT is a concept made possible by technologies
- ⦿ IoT concept allows for collection of massive amount to data
- ⦿ Optimizing for verticals means take out the most of these data...
- ⦿ ... to find correlations, predict trends to...
- ⦿ ... give meaningful information to end-users
- ⦿ It is a huge opportunity to provide low-cost, efficient systems that can be deployed "out-of-the-box"

IOT_2: Unleash the power of IoT data

protocols, analysis, artificial intelligence, machine learning,...

OneTech Booster



Capsule Booster – 2022

Prof. Congduc Pham
<http://www.univ-pau.fr/~cpham>



Horizon 2020
European Union funding
for Research & Innovation



Advanced and disruptive IoT/AI technologies targeting the smallholder community for increased resilience