

CoreML

Comment réconcilier développeur iOS et data scientist ?



❖ LIBERTY RIDER ❖



<https://goo.gl/yjstwe>

<https://goo.gl/yjstwe>



Mathieu Vandeginste (Lead dev iOS)
mathieu@liberty-rider.com



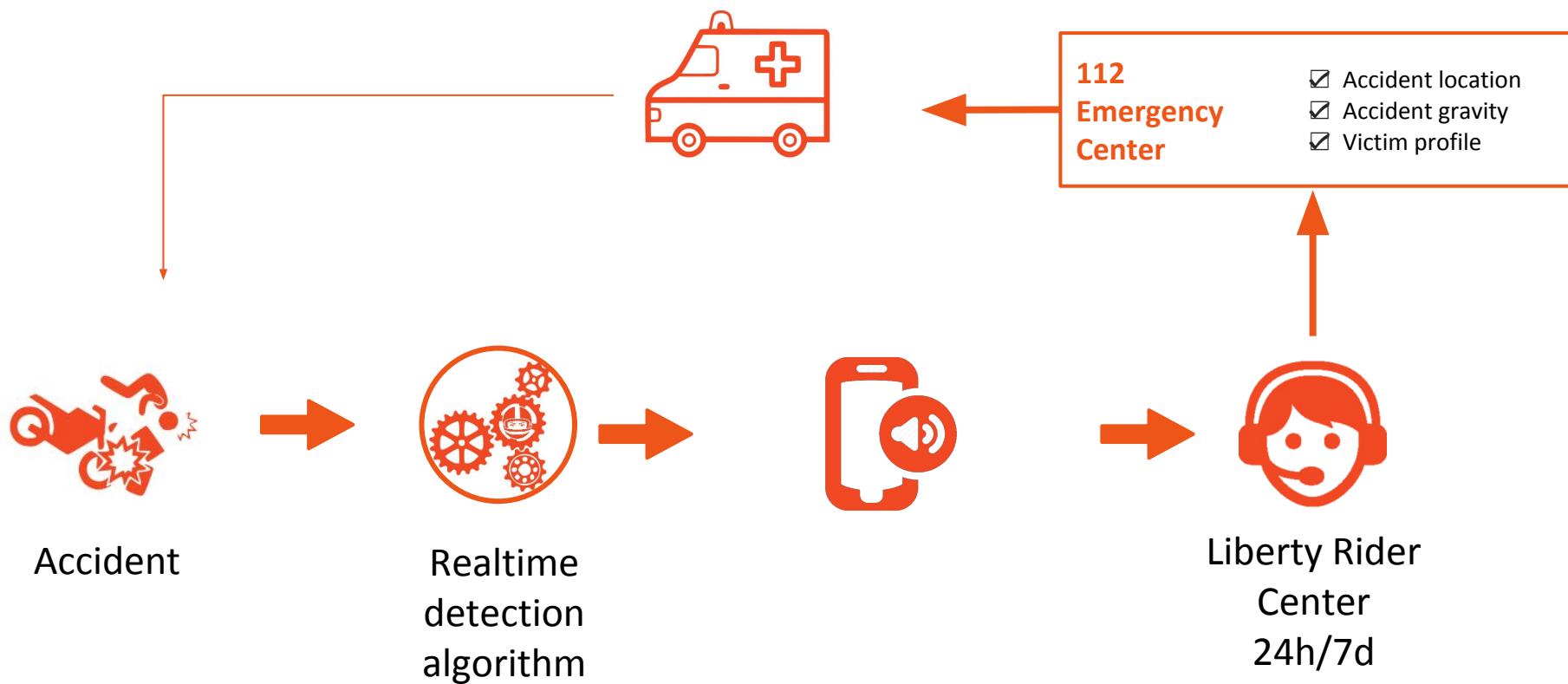
Jérémie Fourmann (R&D lead & co-founder)
jeremie.fourmann@liberty-rider.com
Engineer & Phd

Présentation de Liberty

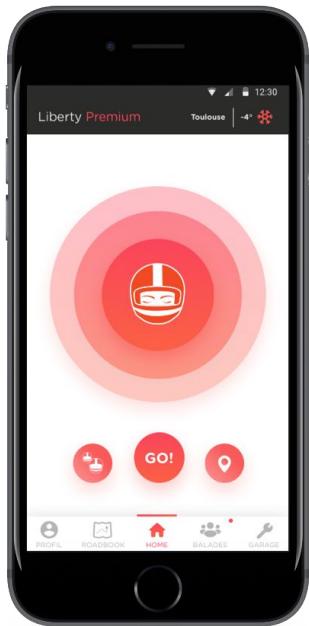


- First motorcyclist protection app
- iOS + Android support
- 160k users in France

Liberty Emergency process



Accident detection Algorithm



Gratuit pour tous les motards !



Different Machine Learning Applications (1/2)

Medical :

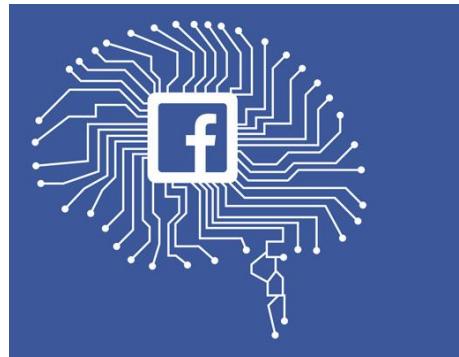


Automotive :



Different Machine Learning Applications (2/2)

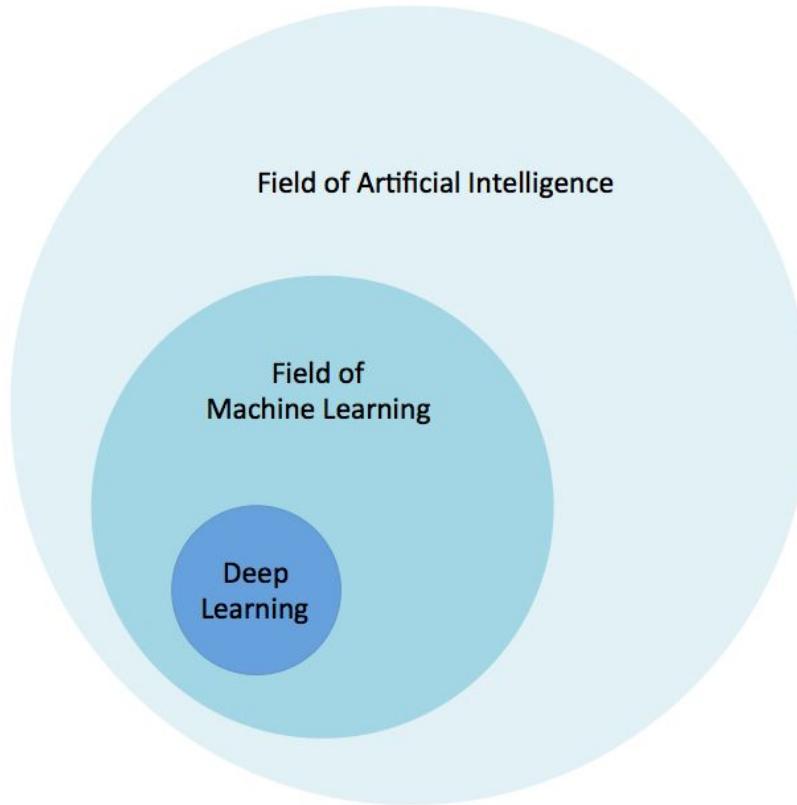
Social network :



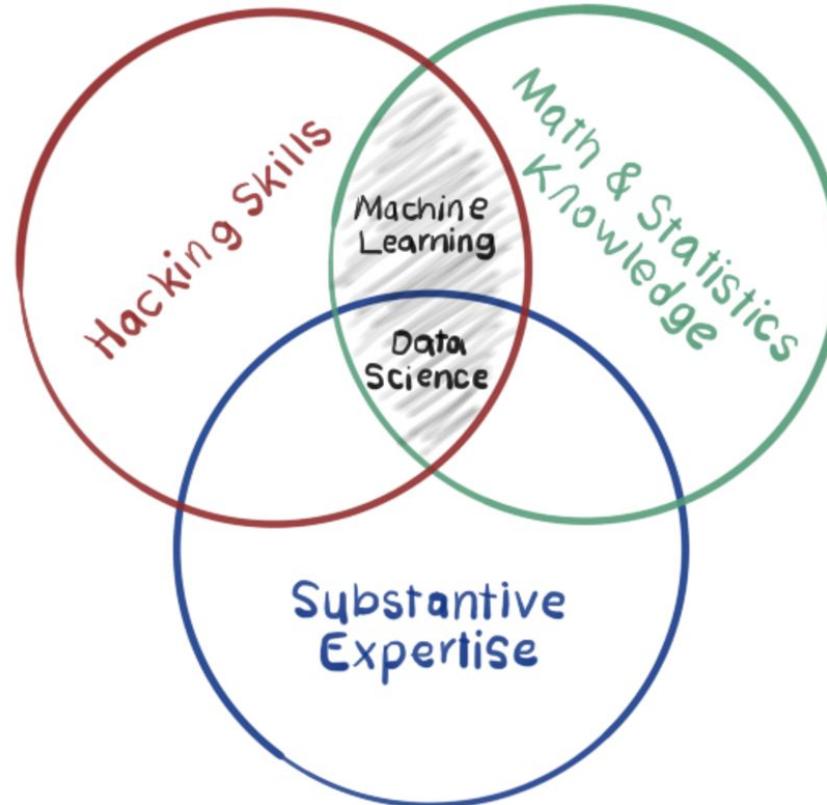
Finance :



AI - Machine - Learning

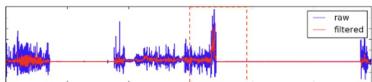


To be a good Machine Learning engineer



Machine learning model

Série temporelle



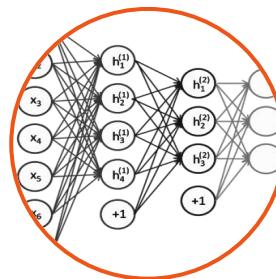
Picture



Text



Input



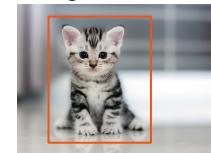
Prediction

neural network (DL)
SVM, Random Forest, Xgboost (ML)

Fall: True - False

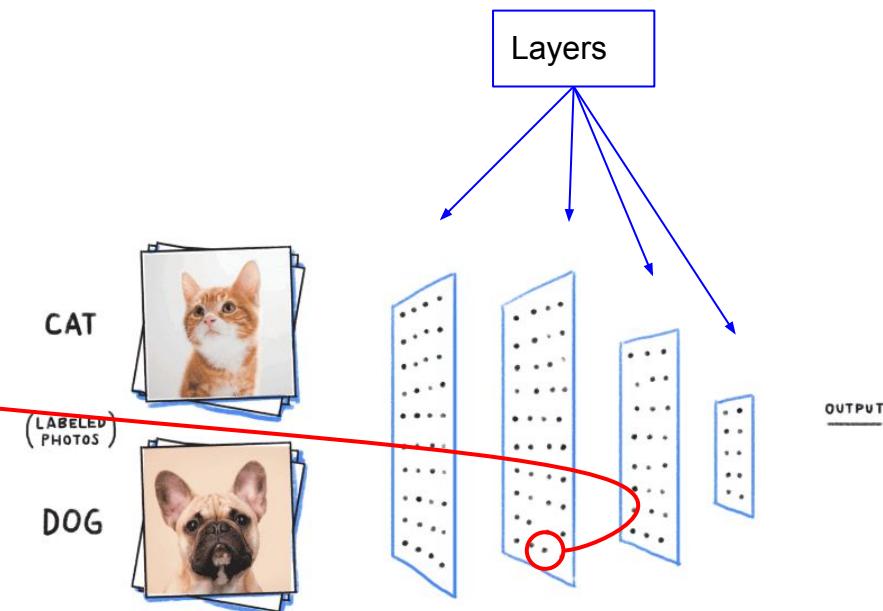
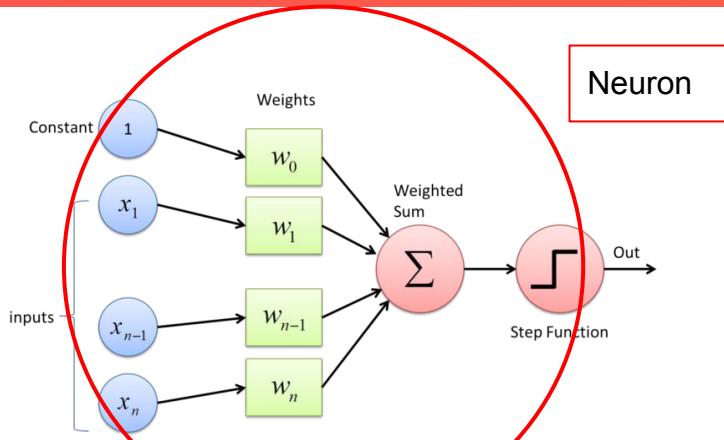


cat: 95%
dog: 5%

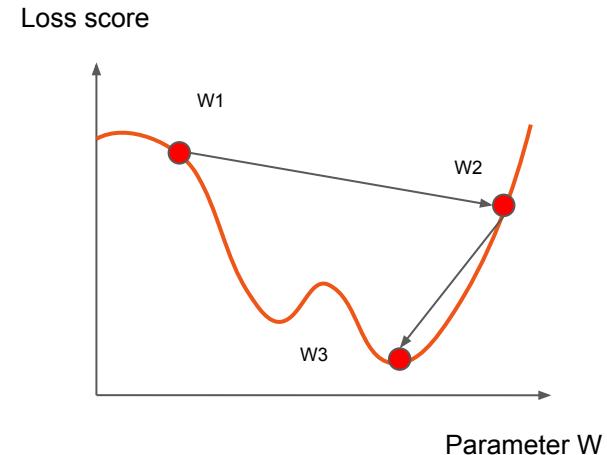
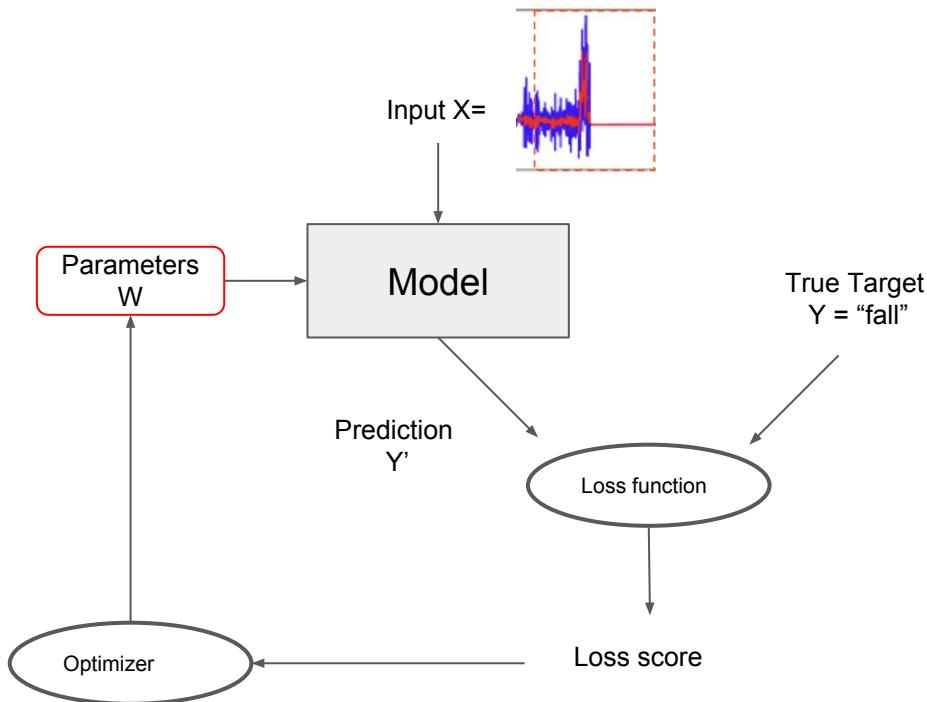


spam, user happiness
level, profile

Basic neural network architecture (Deep learning)



Training process



Entraînement d'un réseau de neurones

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

+

-

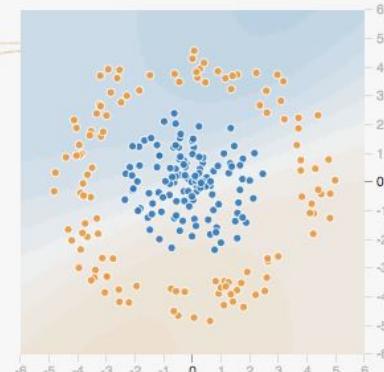
4 neurons

2 neurons

This is the output from one **neuron**. Hover to see it larger.
The outputs are mixed with varying **weights**, shown by the thickness of the lines.

OUTPUT

Test loss 0.518
Training loss 0.513



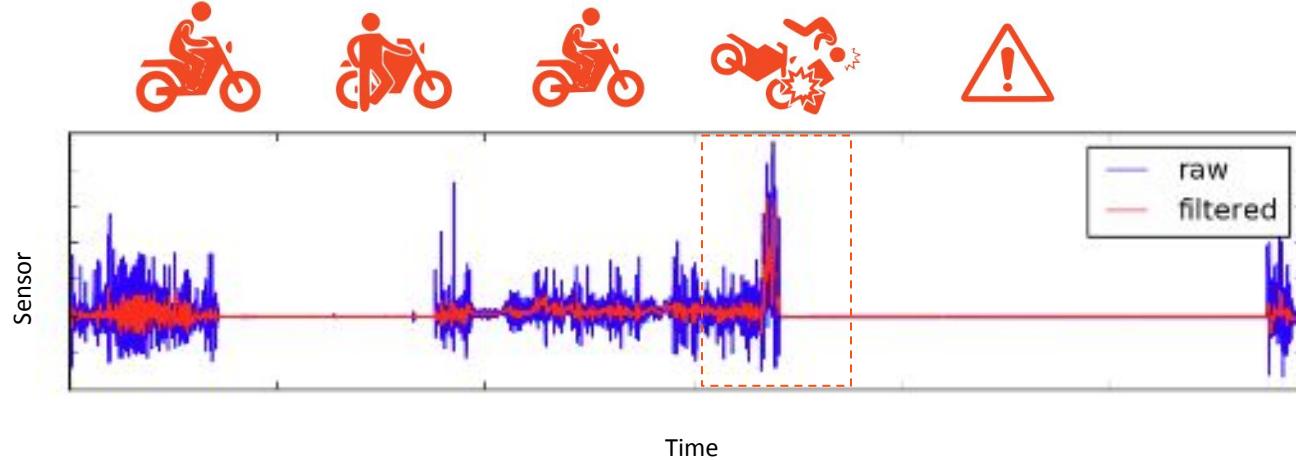
Colors shows data, neuron and weight values.



Show test data

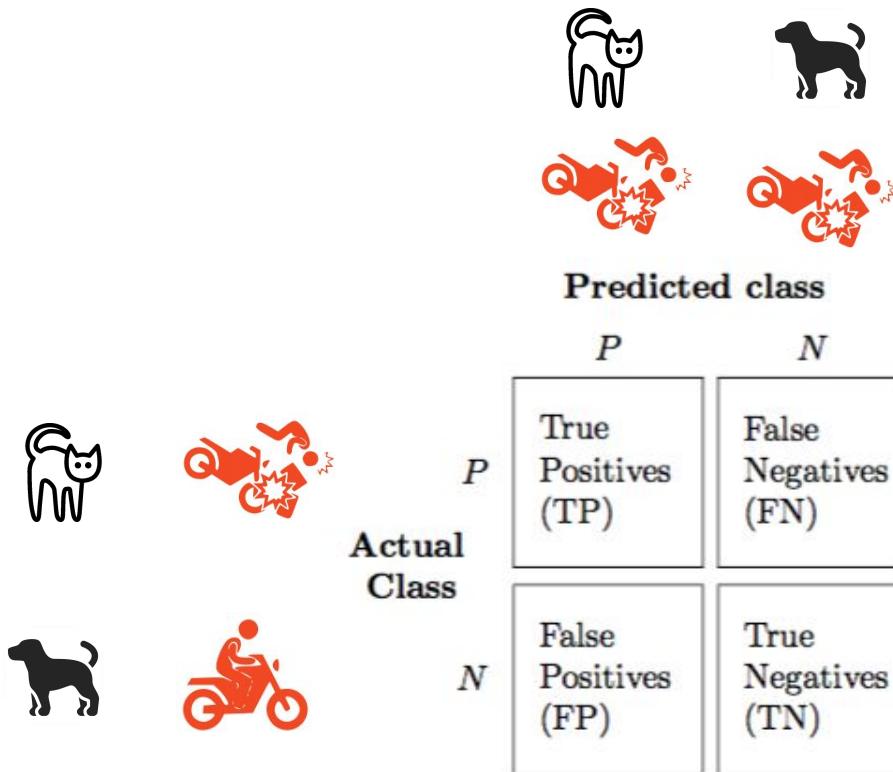
Discretize output

Temporal Domain



Performance indicator : confusion matrix

		Predicted class	
		<i>P</i>	<i>N</i>
Actual Class	<i>P</i>	True Positives (TP)	False Negatives (FN)
	<i>N</i>	False Positives (FP)	True Negatives (TN)

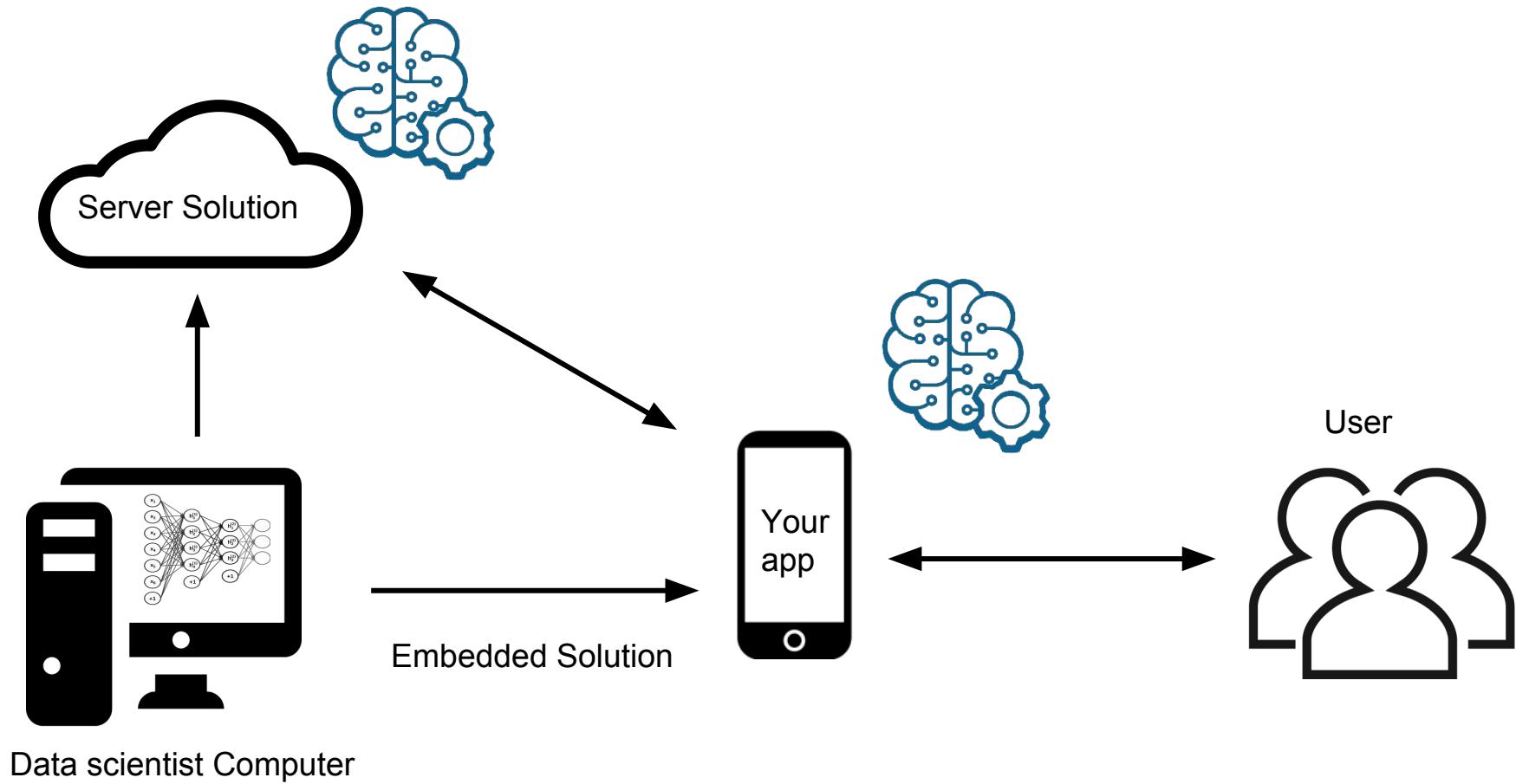


The diagram illustrates a 2x2 confusion matrix for a classification task. The columns represent the predicted classes (P for Positive and N for Negative) and the rows represent the actual classes (P for Positive and N for Negative). The matrix entries are labeled with their respective names: True Positives (TP), False Negatives (FN), False Positives (FP), and True Negatives (TN). The icons for the classes are: a white cat for P (Actual), a black dog for N (Actual), a red motorcycle for P (Predicted), and a red motorcycle with an explosion for N (Predicted).

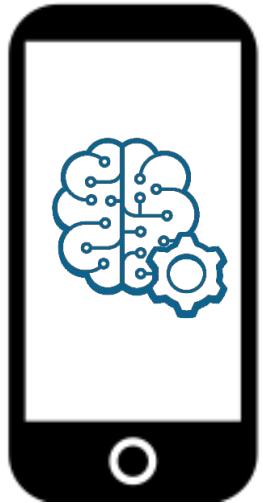
Exemple:

99	1
10	90

Deploy your model for your user



Model embedded in the smartphone



- working without internet
- good prediction performance speed (GPU - CPU)
- scalable
- no server cost



- platform dependent
- difficult to monitor
- update



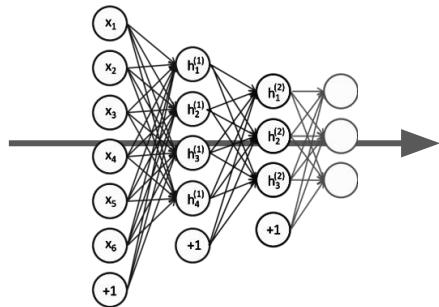
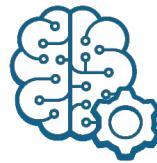
When a Data scientist interact with an iOS Developer

Data scientist



“Voici le nouveau modèle”

ML Model

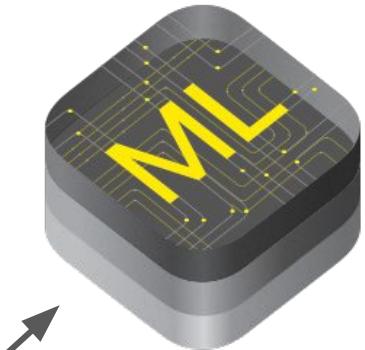
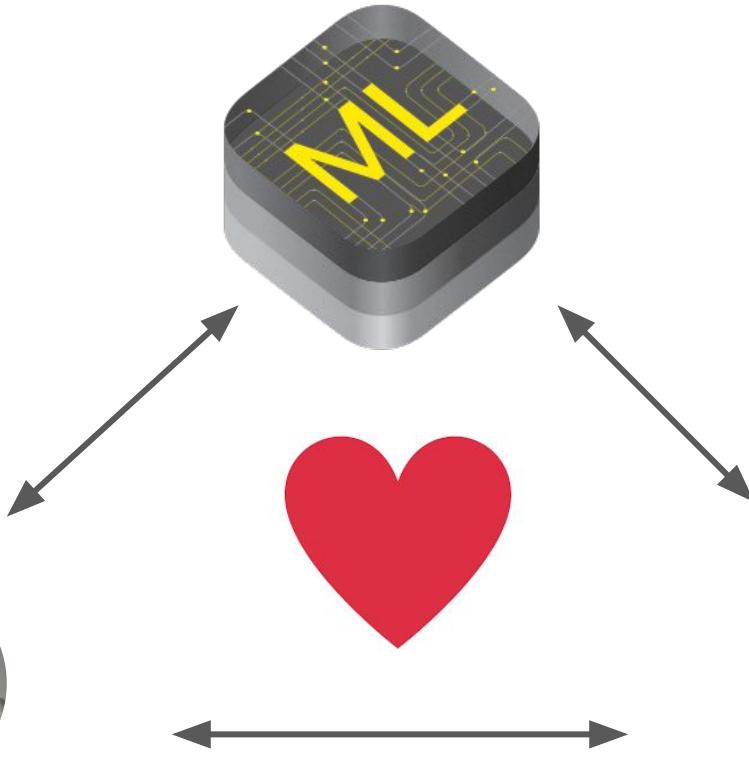


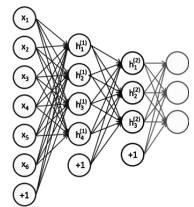
Developer



“WTF”

CoreML: make data scientist & ios dev love again

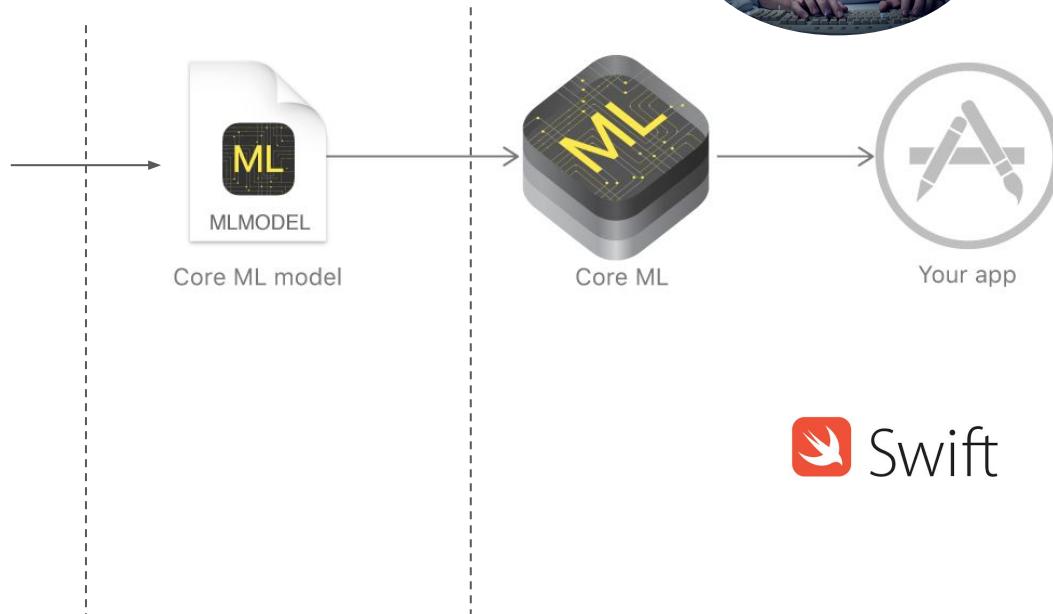




python™

TensorFlow™

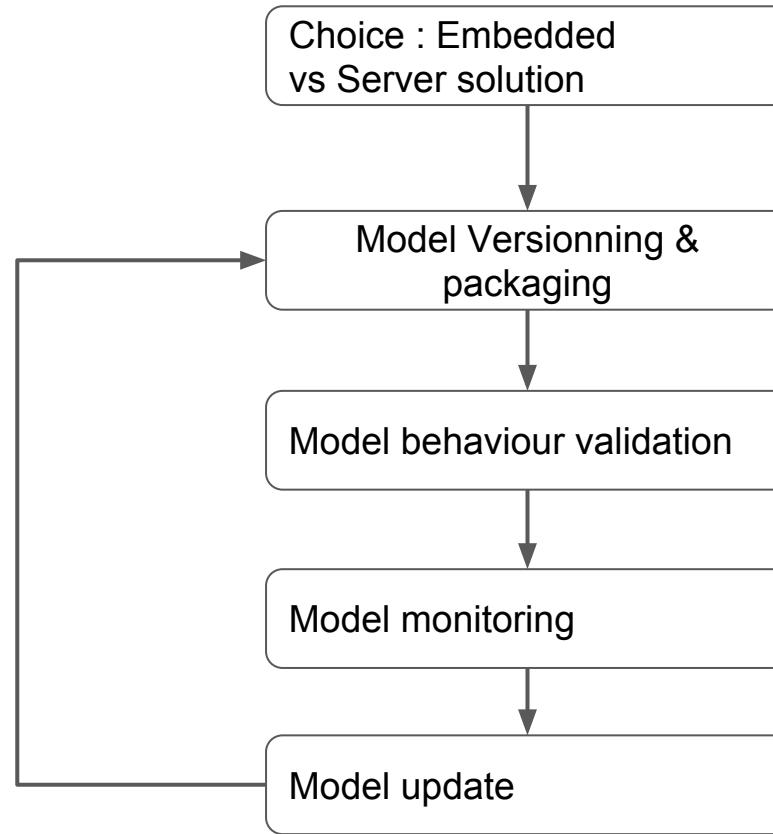
scikit-learn Keras



Swift



ML model in production guideline



Présentation



Dataset presentation



12500 : cat pictures

12500 : dog pictures

Processing: resize + scale

1000 cats
1000 dogs

400 cats
400 dogs

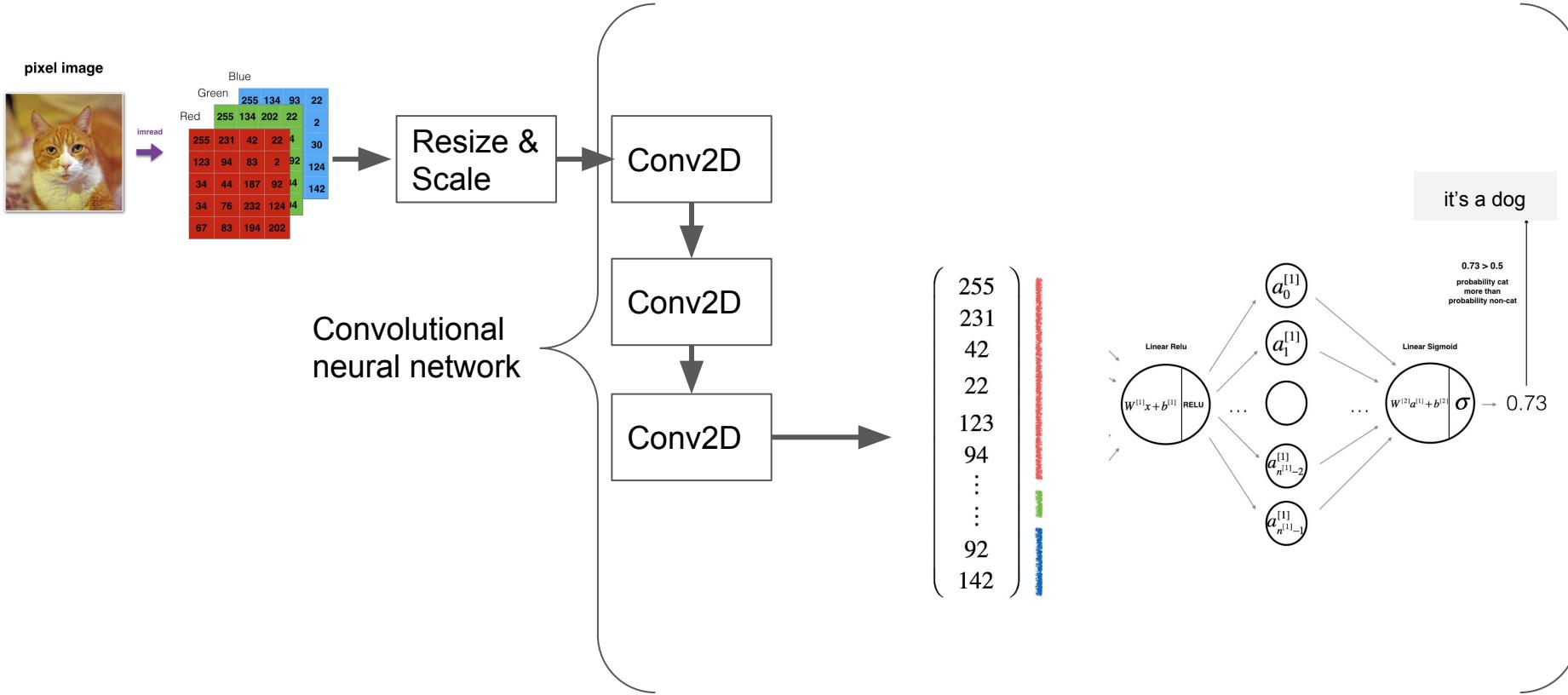
400 cats
400 dogs

Training

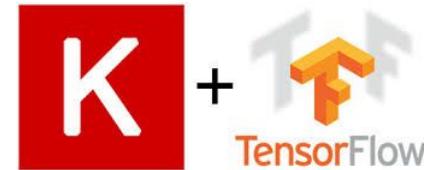
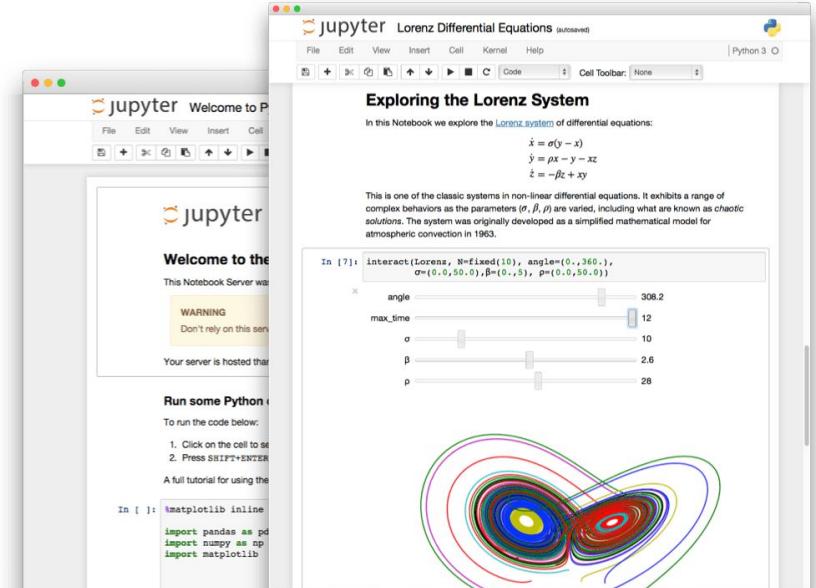
Validation

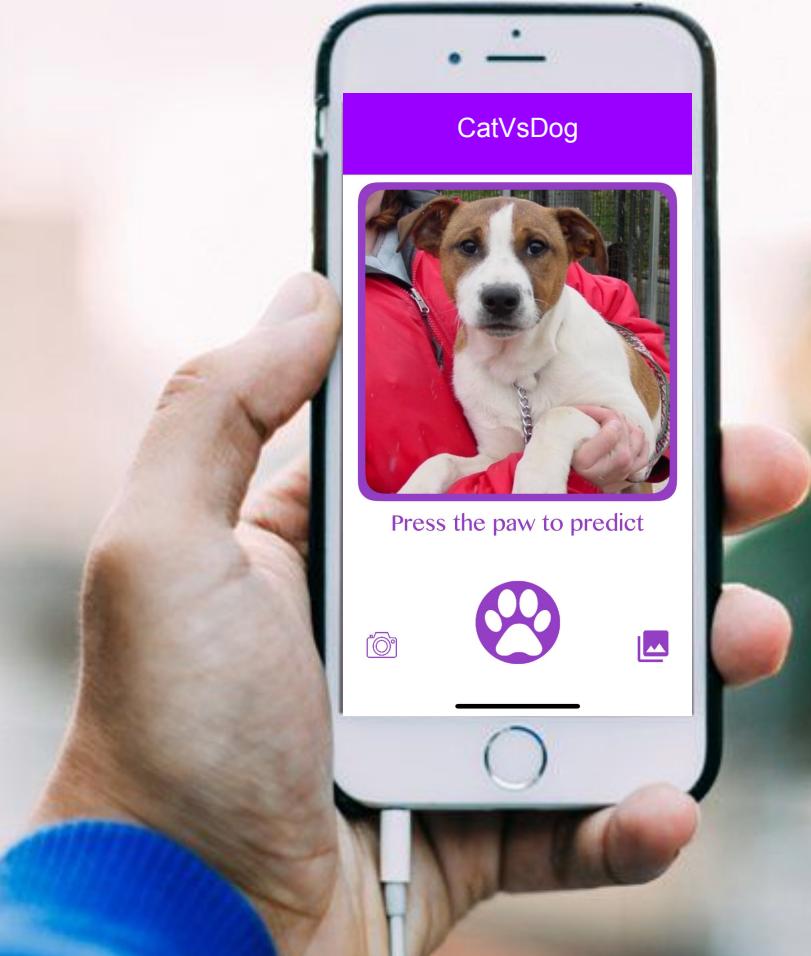
Test

Network architecture presentation



Python datascience environnement





Available on the
App Store

LIVE  CODING



✓ LIBERTY RIDER ✓

ANGE GARDIEN DES MOTARDS

We are Hiring

<http://www.liberty-rider.com/careers/>



contact us : jobs@liberty-rider.com

Annexes



NewFeature



data scientist: C'est bon j'ai un super modèle, tu peux le coder ?



iOS dev: ok fait moi une PR sur le projet



data scientist: une quoi ? non j'ai juste les poids à te passer lors de la mise en prod, peux-tu implémenter un réseau de neurone [5 2 1], regarde sur wikipédia pour plus d'info



iOS dev: ok c'est codé, j'ai besoin des poids pour l'initialisation



data scientist: je te les ai envoyés par slack, c'est dans un csv



iOS dev: ok c'est en prod



Product owner: la nouvelle feature ne marche pas du tout , c'est quoi ce bordel !!!

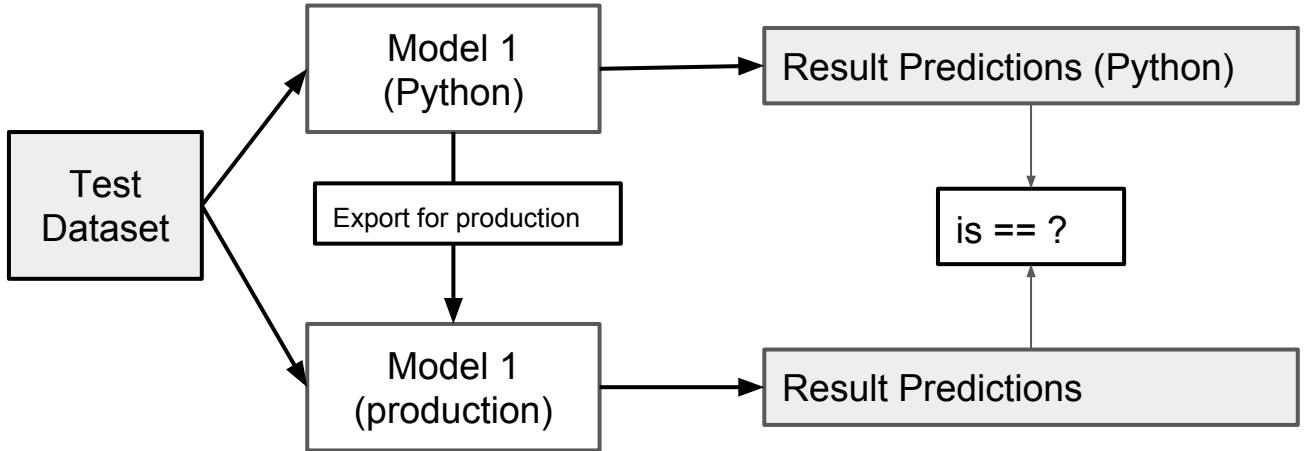


data scientist: c'est normal , j'ai fait une maj des poids, tu a mis quel poids dans le modèle ?



iOS dev: j'en sais rien , ce que tu m'as slacké ?

Model Behaviour Validation



Unit & integration Test

Ressources:

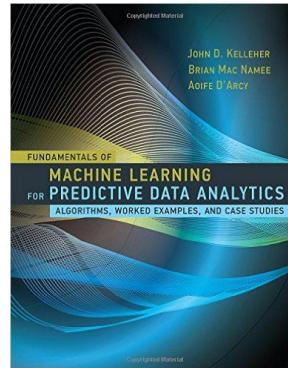
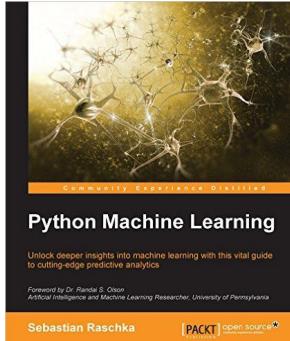
Blog :

www.machinelearningmastery.com/
www.pyimagesearch.com/

Online courses :

www.coursera.org/learn/machine-learning
www.coursera.org/learn/deep-learning
www.fast.ai

Book :



Make data science with people:

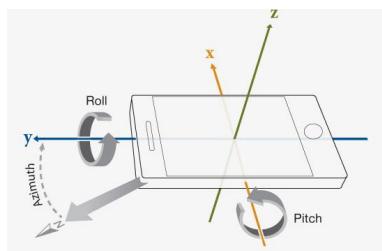
Kaggle:



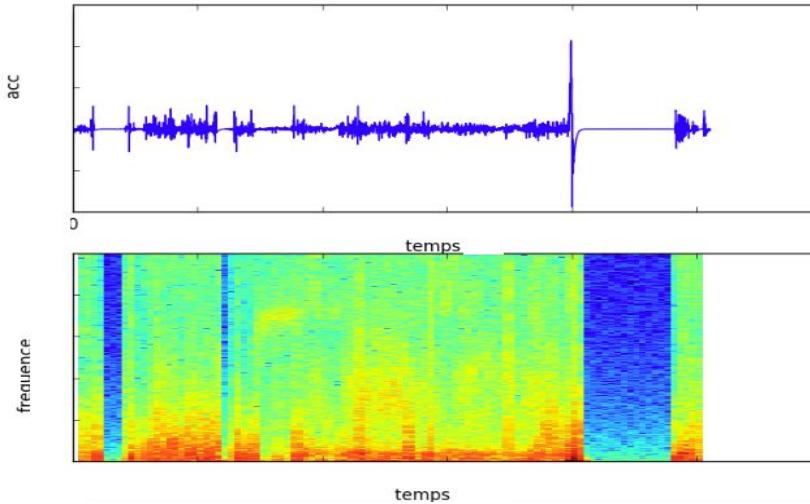
Toulouse Data Science:



1 sensor: the accelerometer

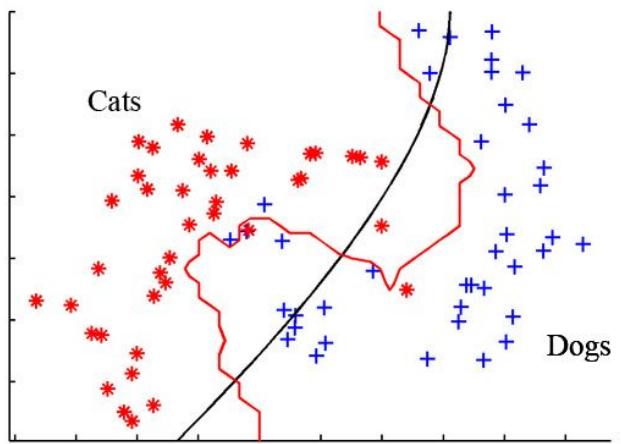


- low battery consumption
- rich source of information in temporal and frequency domain
- every smartphone has an accelerometer

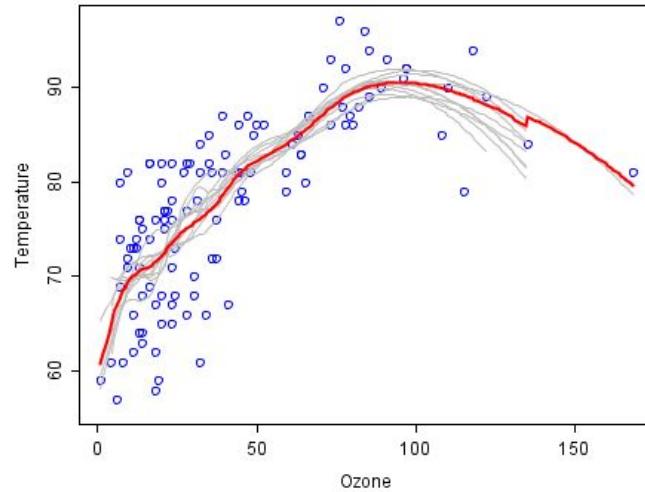


Supervised machine learning

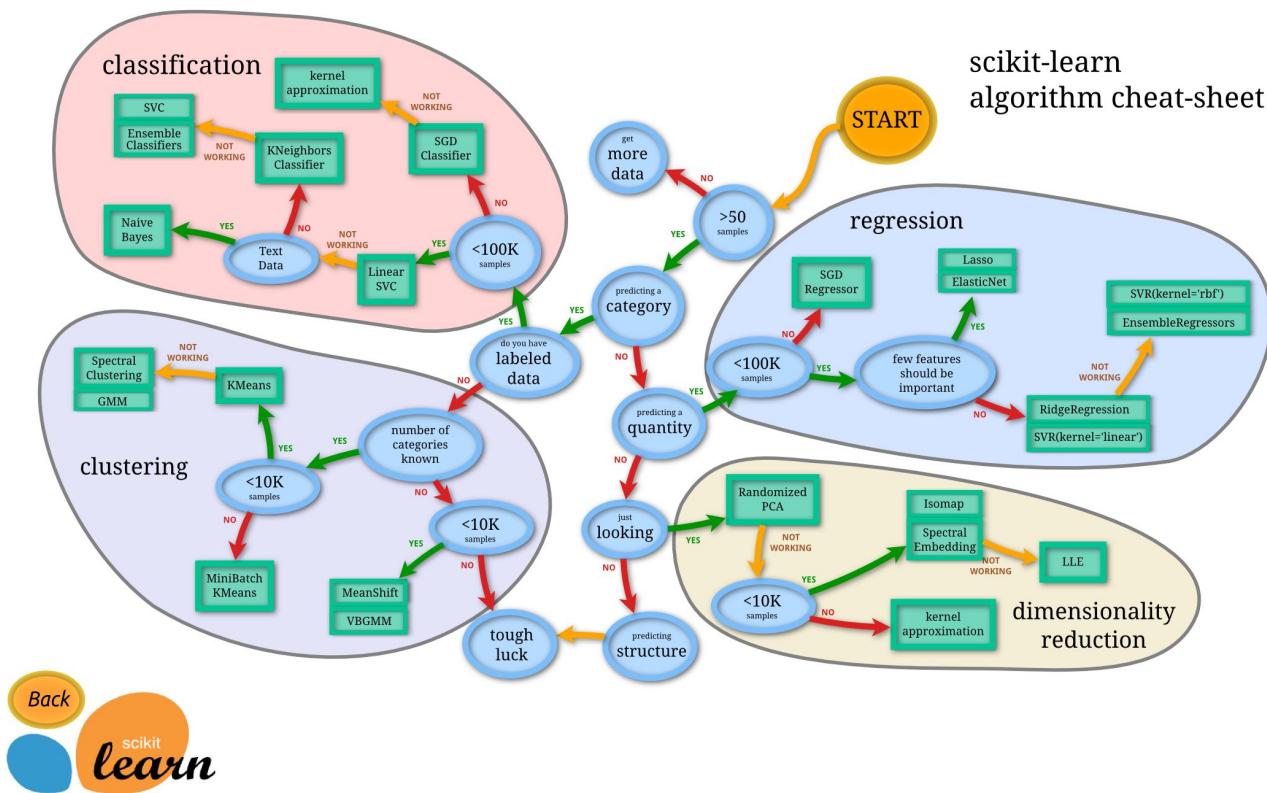
Classification



Regression



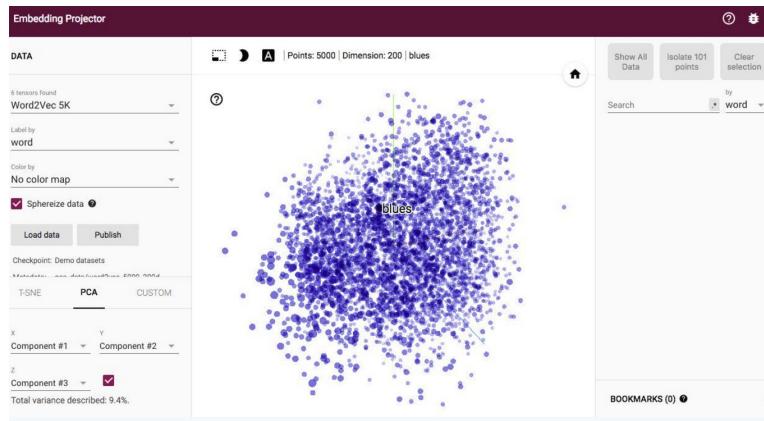
Algorithm cheat-sheet



Our analysing tool:

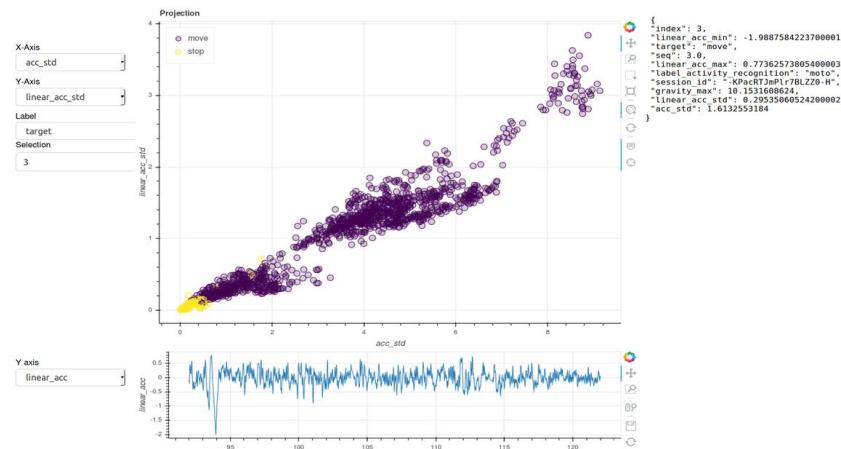
Tensor flow projector (Google):

- dimension reduction (PCA) visualisation
- meta-data analysis



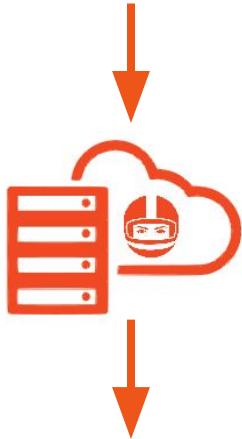
Projector app (Liberty-Rider):

- 2D-features projections
- time series- features visualisation



Accident Detection API Example

```
curl -X POST -d '{feature_1:0.125, feature_2:21558, ..., feature_n=122}'  
-H 'Content-Type: application/json' -H 'Authorization: Simple YOUR_API_KEY'  
https://api.liberty-science.com/accident-prediction/v3
```



```
HTTP 200 OK  
'{accident:true}'
```