The Deep Learning Codelab

Jeff Abrahamson Hugo Mougard

19 octobre 2017

DevFest Nantes 2017

Machine Learning

Definition

- Supervised
- Unsupervised
- Reinforcement

What is ML ?

Machine learning is not magic

What is ML ?

Machine learning is mathematics

What is ML ?

Mostly, it's these maths :

- Probability
- Statistics
- Linear algebra
- Optimisation theory
- Differential calculus

What is Statistics

1. Identify a question or problem.
2. Collect relevant data on the topic.
3. Analyze the data.
4. Form a conclusion.

What is Statistics

1. Identify a question or problem.
2. Collect relevant data on the topic.
3. Analyze the data.
4. Form a conclusion.

Sadly, sometimes people forget 1.

What is Statistics

1. Identify a question or problem.
2. Collect relevant data on the topic.
3. Analyze the data.
4. Form a conclusion.

Statistics is about making 2–4 efficient, rigorous, and meaningful.

OpenIntro Statistics, 2nd edition, D. Diez, C. Barr, M. Çetinkaya-Rundel, 2013.

What is data science ?

(Exercise : Is this the same question as the last slide ?)

1. Define the question of interest
2. Get the data
3. Clean the data
4. Explore the data
5. Fit statistical models
6. Communicate the results
7. Make your analysis reproducible

What is data science ?

(Exercise : Is this the same question as the last slide ?)

1. Define the question of interest
2. Get the data
3. Clean the data
4. Explore the data
5. Fit statistical models
6. Communicate the results
7. Make your analysis reproducible

What the public thinks.

What is data science ?

(Exercise : Is this the same question as the last slide ?)

1. Define the question of interest
2. Get the data
3. Clean the data
4. Explore the data
5. Fit statistical models
6. Communicate the results
7. Make your analysis reproducible

Where we spend most of our time.

What is data science ?

(Exercise : Is this the same question as the last slide ?)

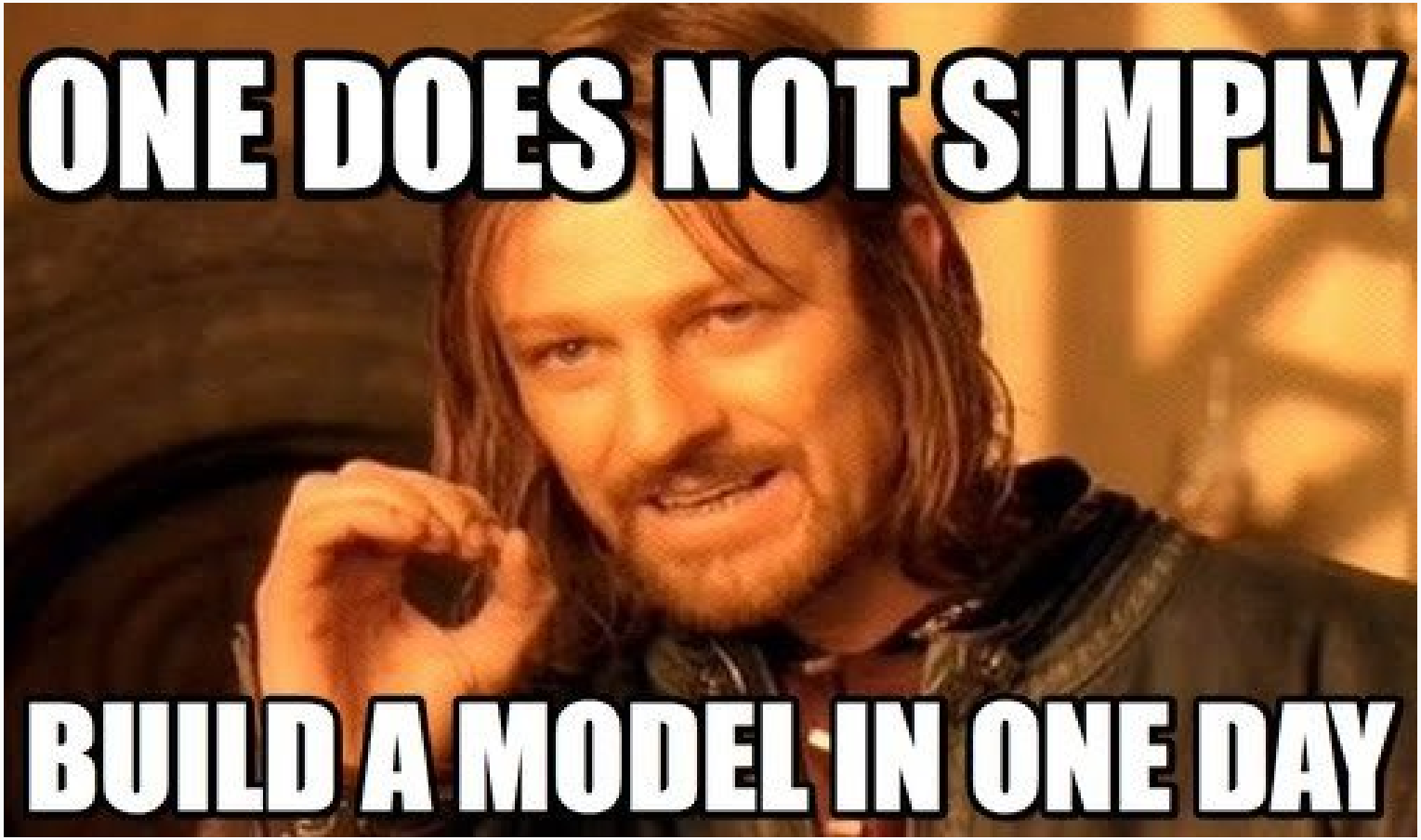
1. Define the question of interest
2. Get the data
3. Clean the data
4. Explore the data
5. Fit statistical models
6. Communicate the results
7. Make your analysis reproducible

The easiest part to forget.

What is data science ?

http://simplystatistics.org/2015/03/17/data-science-done-well-looks-easy-and-that-is-a-big-problem-for-data-scientists/

What is data science ?



Representation

Typically a vector space

Features are dimensions

Features

Feature extraction

Feature engineering, synthetic features

Feature Engineering

1. Brainstorm
2. Pick some
3. Make them
4. Evaluate
5. Repeat

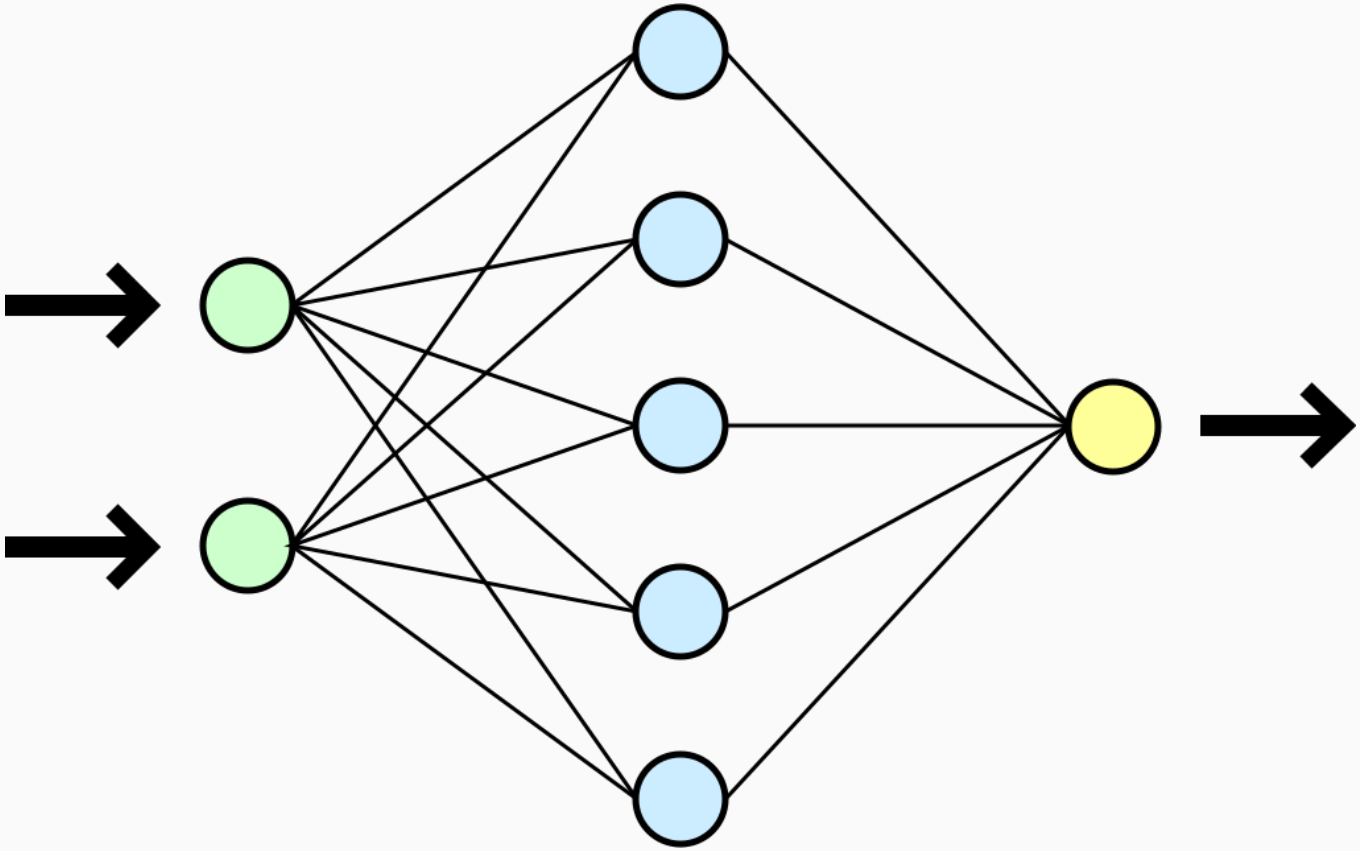
Artificial Neural Networks

Neural Networks

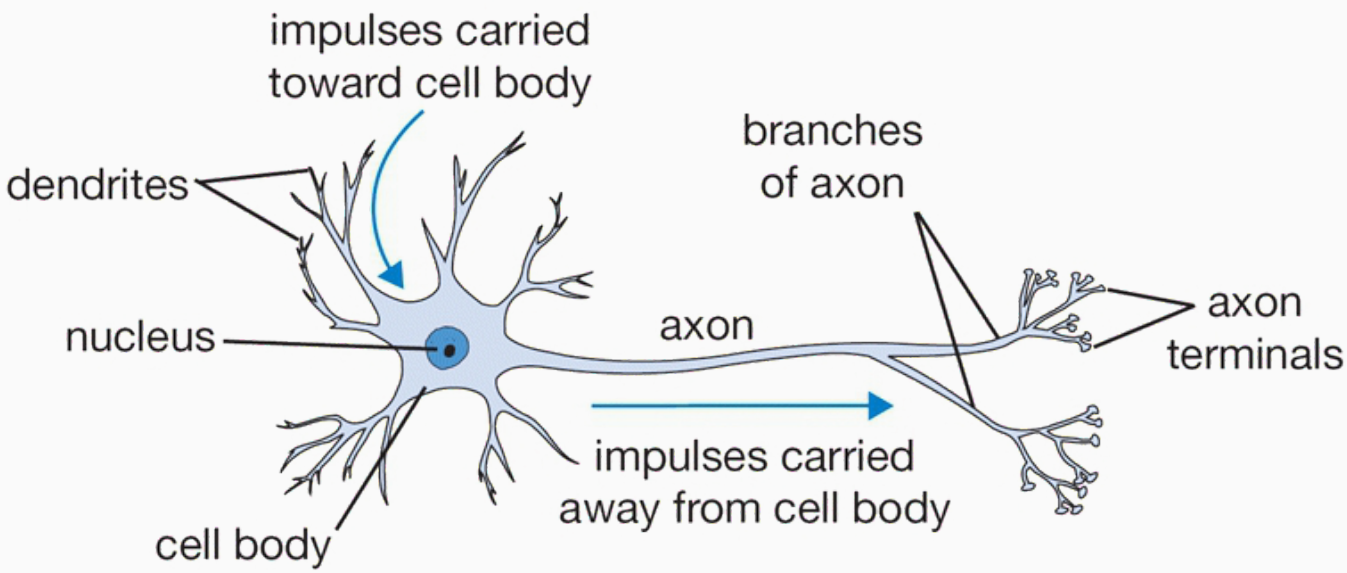
Inspired by brain :

- 1 neuron = 1 simple computation
- 10M neurons = 10M simple computations
- 10M combined neurons = 10M simple computations combined = 1 interesting computation

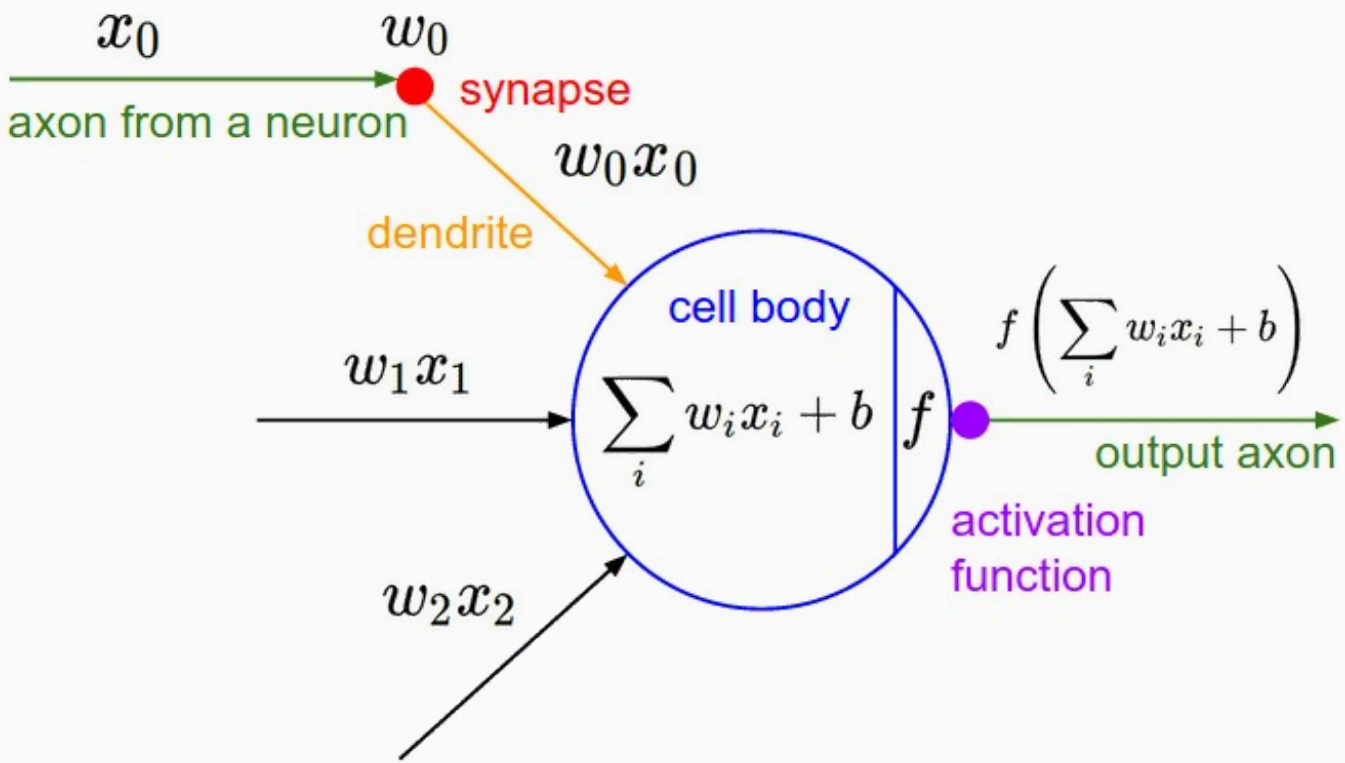
Neural Networks — Neural Network Example



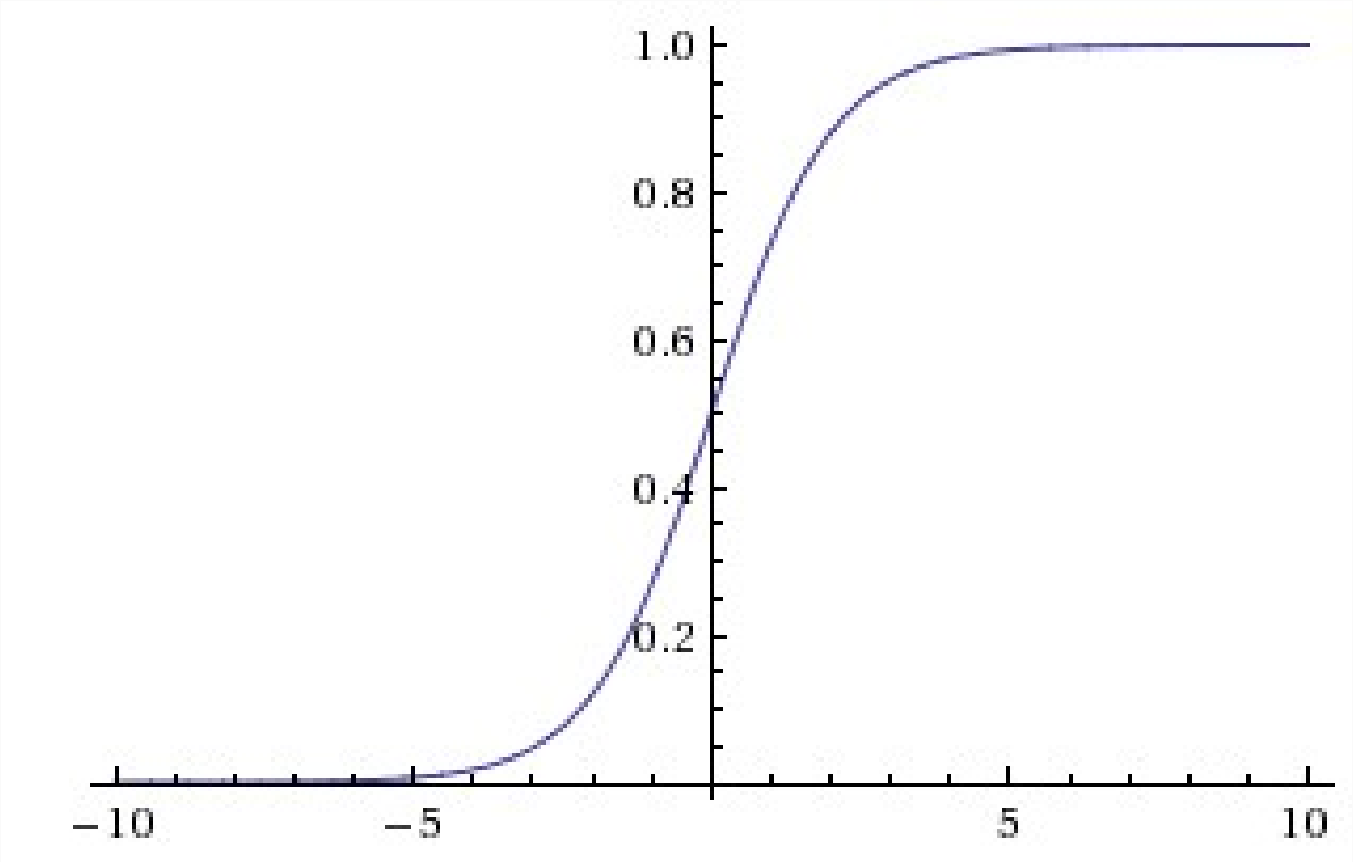
Neural Networks — Biological Neuron



Neural Networks — Artificial Neuron



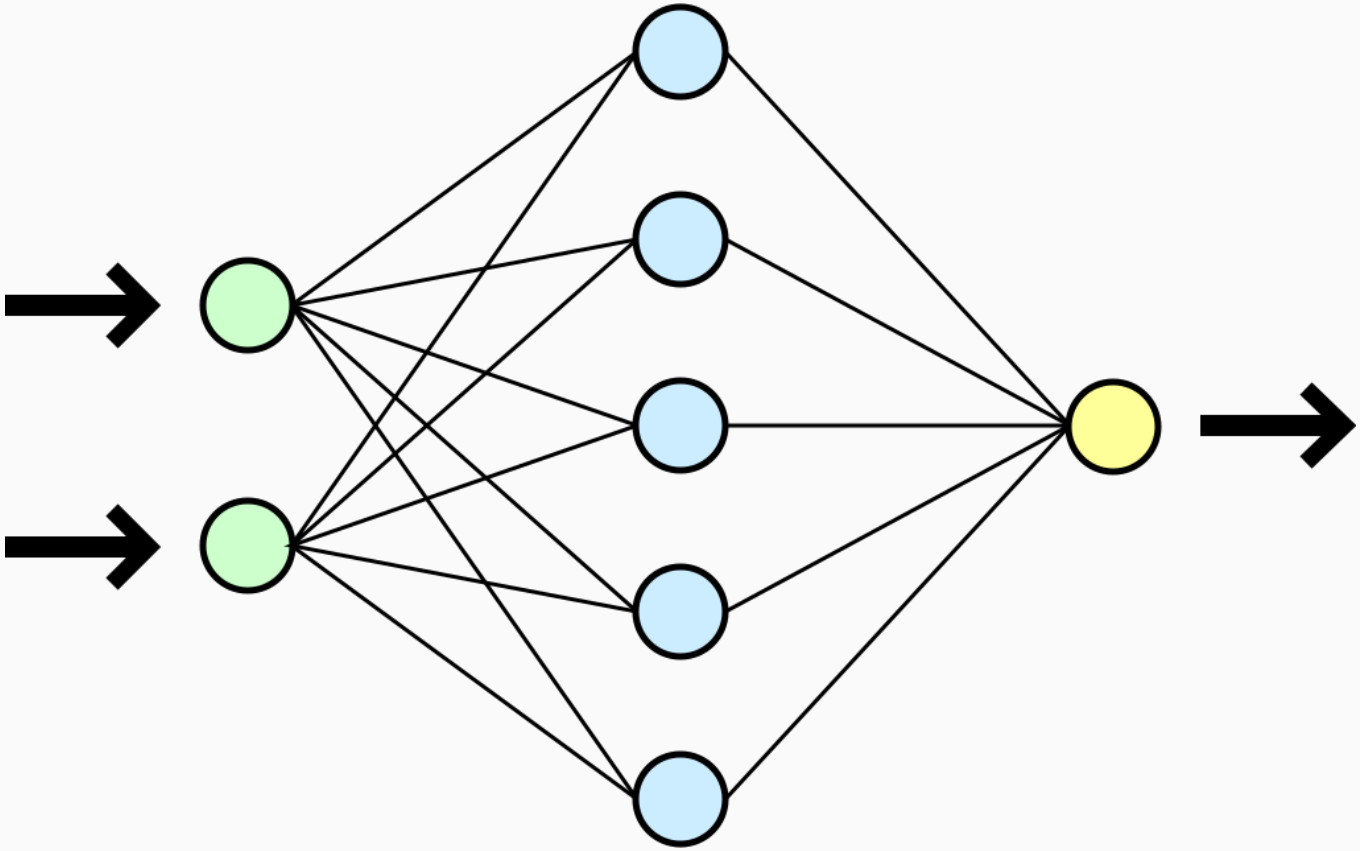
Neural Networks — Activation Function



Neural Networks — So What's a Neuron Again ?

- Neuron = Weighted averaged of its input + activation function
- Goal of learning = learn good weights.

Neural Networks — Neural Network Example Comeback



Neural Networks — How to know if they're good

- Find some good examples
- Compare network output and good examples (Loss function).

Neural Networks — How to Learn Weights ?

1. A Neural Network loss is a function
2. You actually know how to minimize functions (derivatives)
3. ...
4. Profit

Convolutionnal Neural Networks

Quel besoin ?

Réseau standard + images = Too Much Information.

Keras

Start with a plan

1. Load Data.
2. Define Model.
3. Compile Model.
4. Fit Model.
5. Evaluate Model.
6. Tie It All Together.

MNIST

Hey I'm a frame

Hello, world !

Infra

Hey I'm a frame

Hello, world !

Model

Hey I'm a frame

Hello, world !

Resources

Resources

Hello, world !

Q & A