AI for HEP

David Rousseau, IJCLab-Orsay

david.rousseau@ijclab.in2p3.fr

CHACAL, Johannesburg, Jan 2024

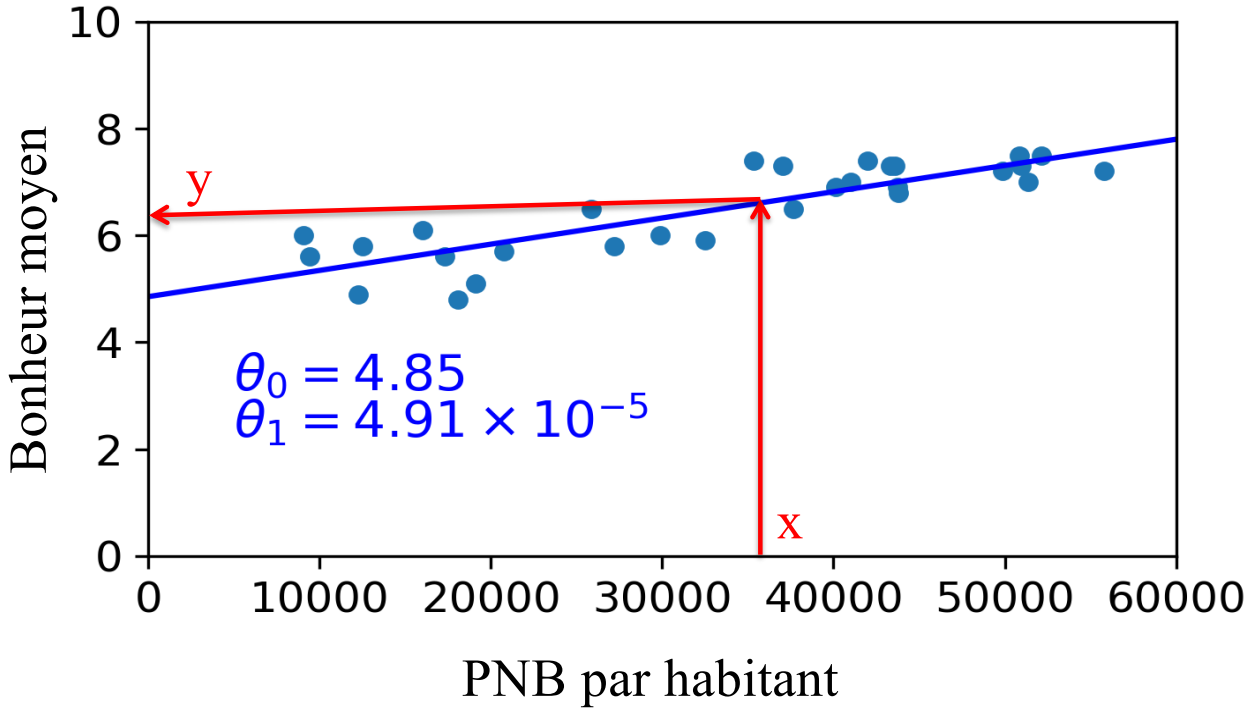
Outline

- Broad overview of AI for science : now
- Basics of AI for physics + BDT tutorial : now and this PM
- Neural Networks course and tutorial : friday
- Advance ML for HEP : saturday

AI for science

Linear Regression

Boskovic, Legendre, Laplace, Gauss



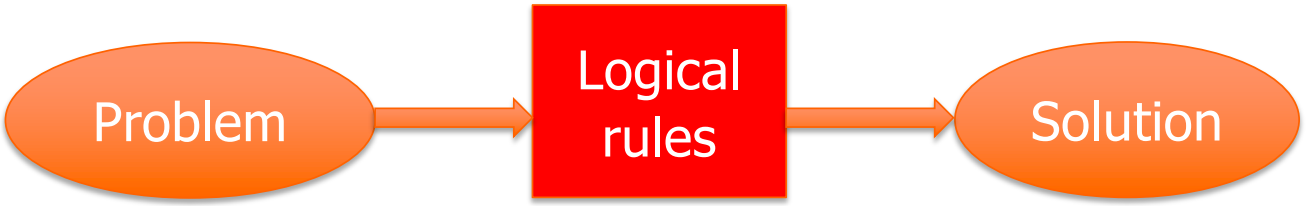
$y = f(x)$

Given x, we want y ➔ how to build f ?

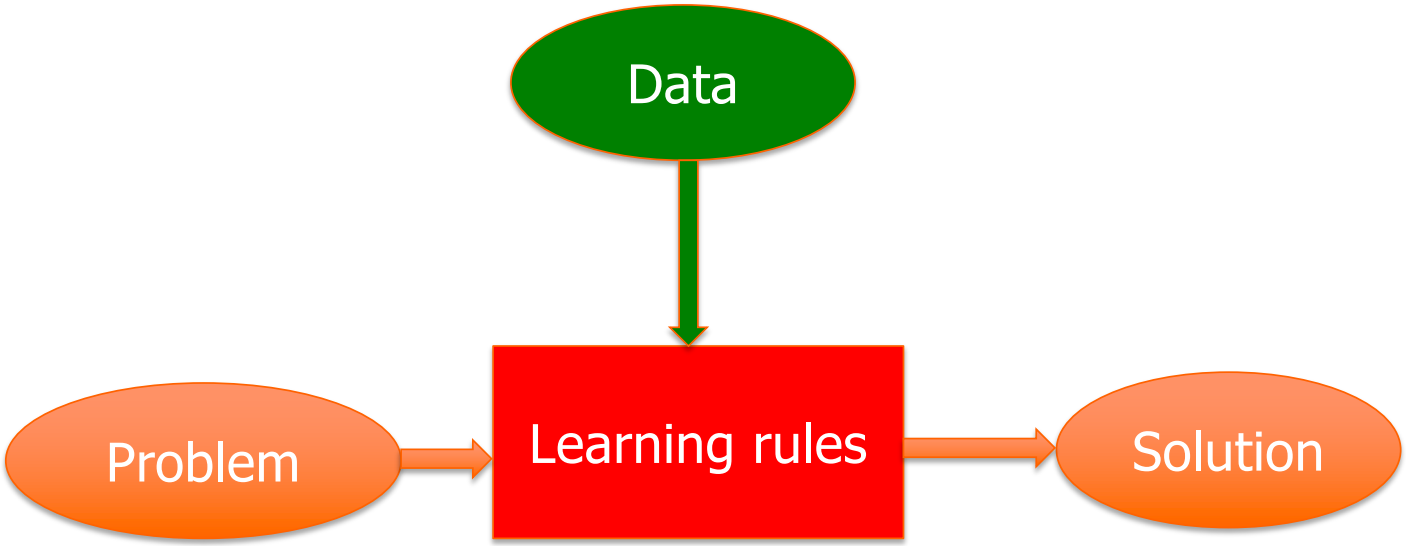
x f y

- Written text ➔ text
- Image ➔ cat or dog ? Mom ?
- « Comment ça va ? » ➔ «Unjani? »
- Speach ➔ text
- Facebook account. ➔ ads
- Chess diagram ➔ next move
- Camera+GPS ➔ steering wheel action
- « write an essay on AI and science »➔ essay

Traditional computing



Machine Learning

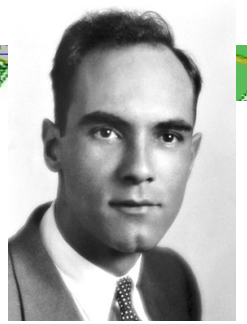


Learning = = optimising algorithm internal parameters: n=2 à 10^{12}

Scientific inference

Discovery of the positron

- Cloud chamber



Carl Anderson
Nobel 1936

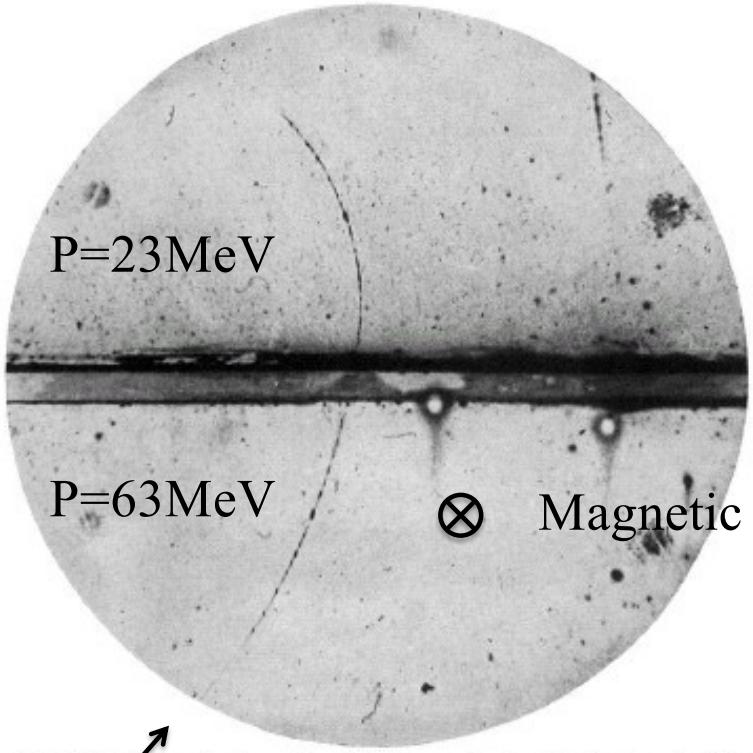
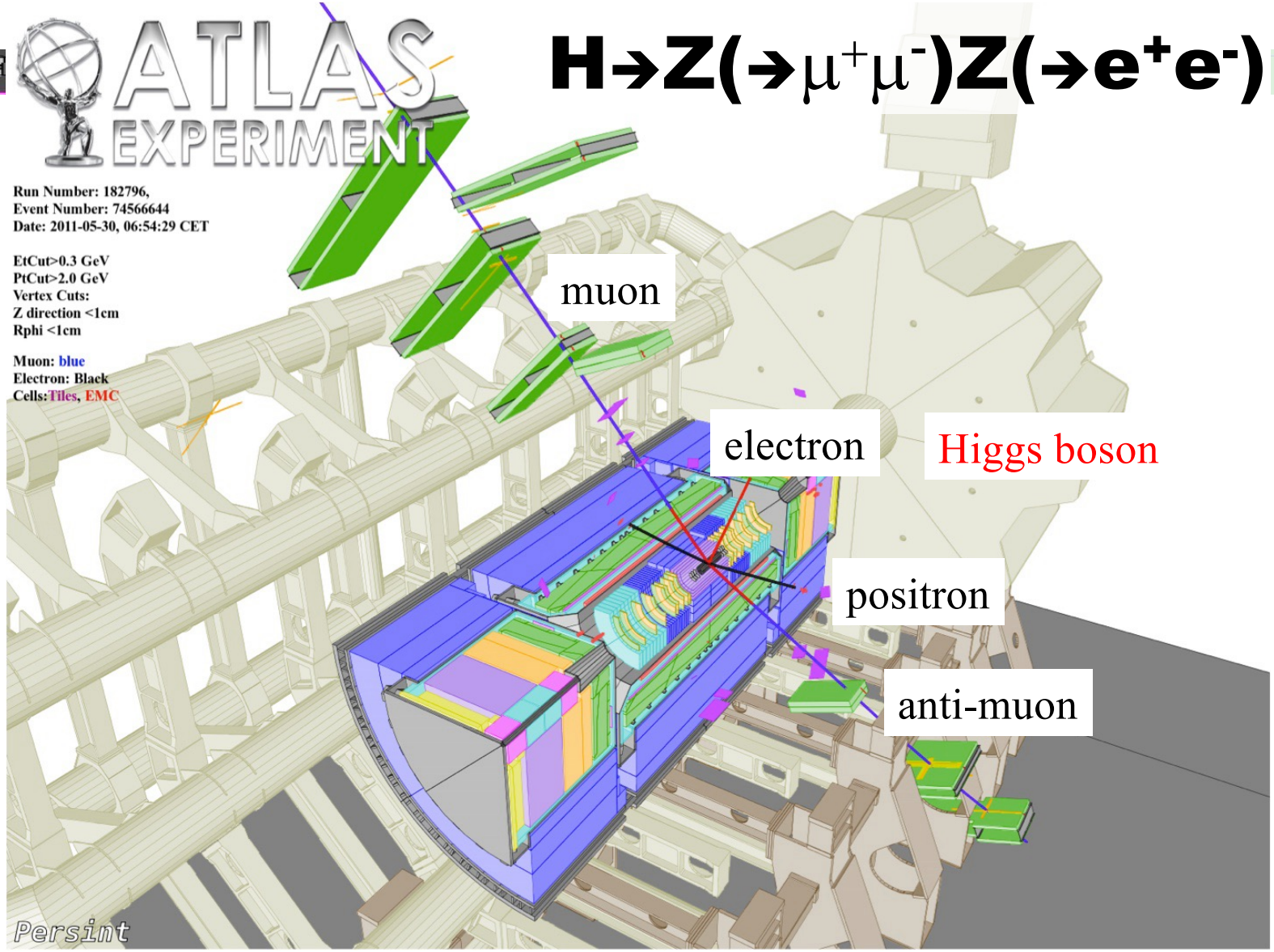


FIG. 1. A 63 million volt positron ($H\rho = 2.1 \times 10^5$ gauss-cm) passing through a 6 mm lead plate and emerging as a 23 million volt positron ($H\rho = 7.5 \times 10^4$ gauss-cm). The length of this latter path is at least two times greater than the possible length of a proton path of this curvature.

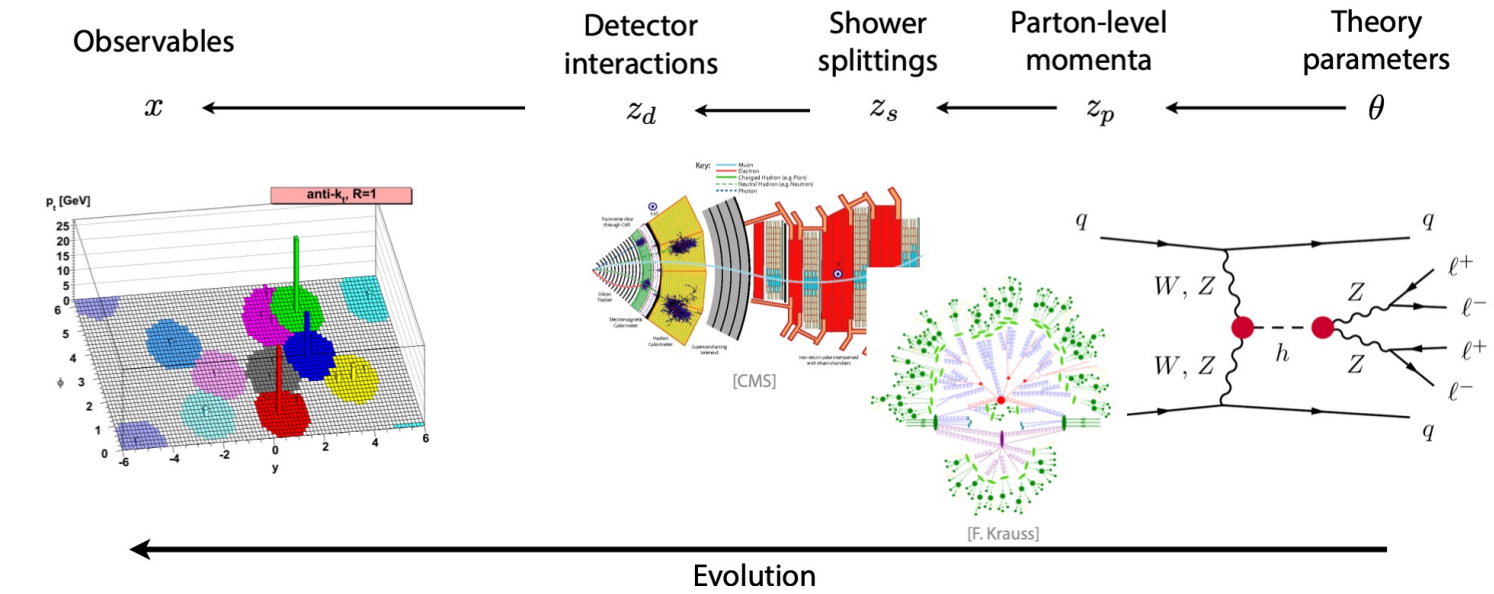
Inferences

Einstein: $E=mc^2$

$H \rightarrow Z(\rightarrow \mu^+\mu^-)Z(\rightarrow e^+e^-)$



Modeling and Inference

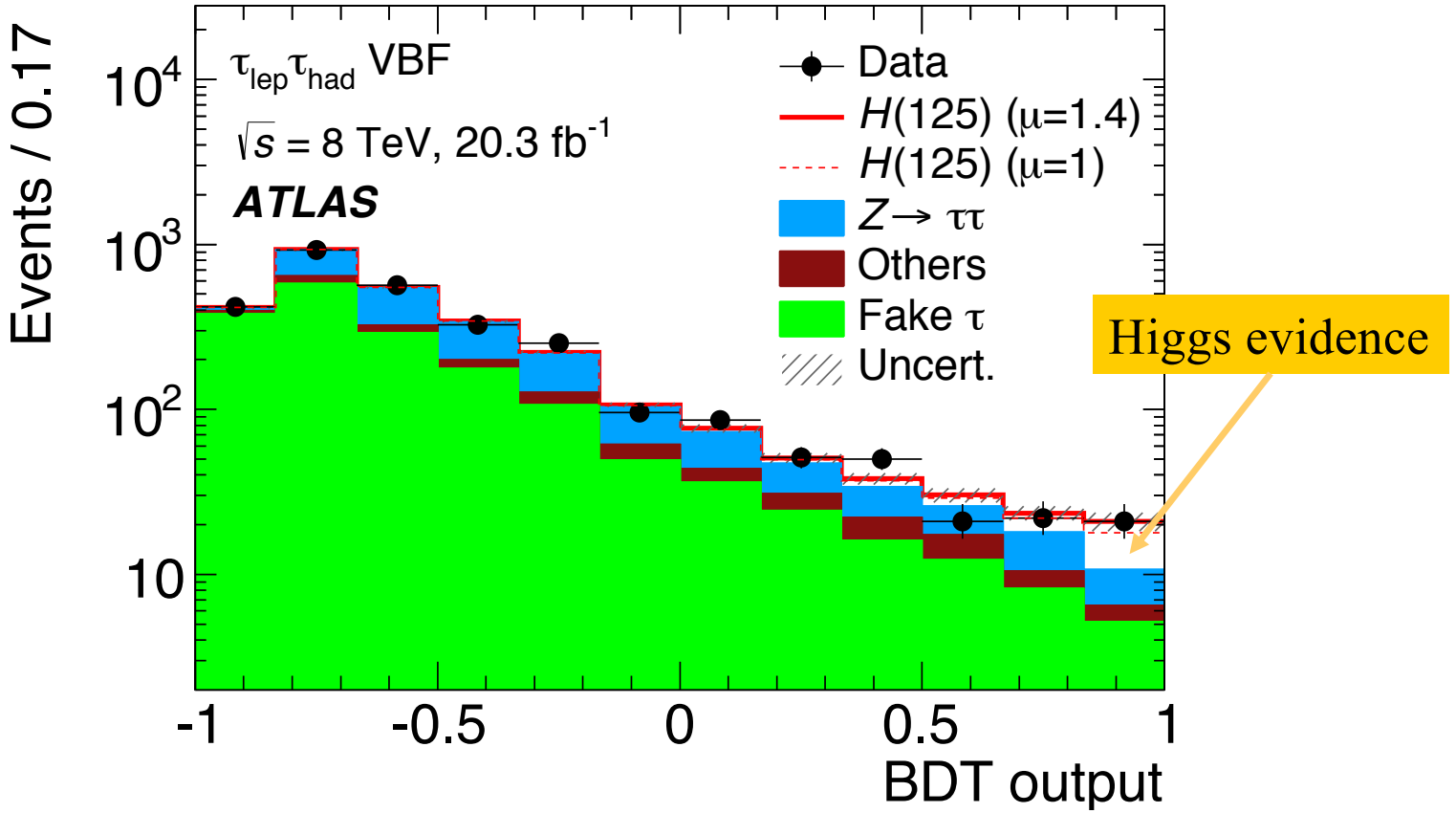


Inference

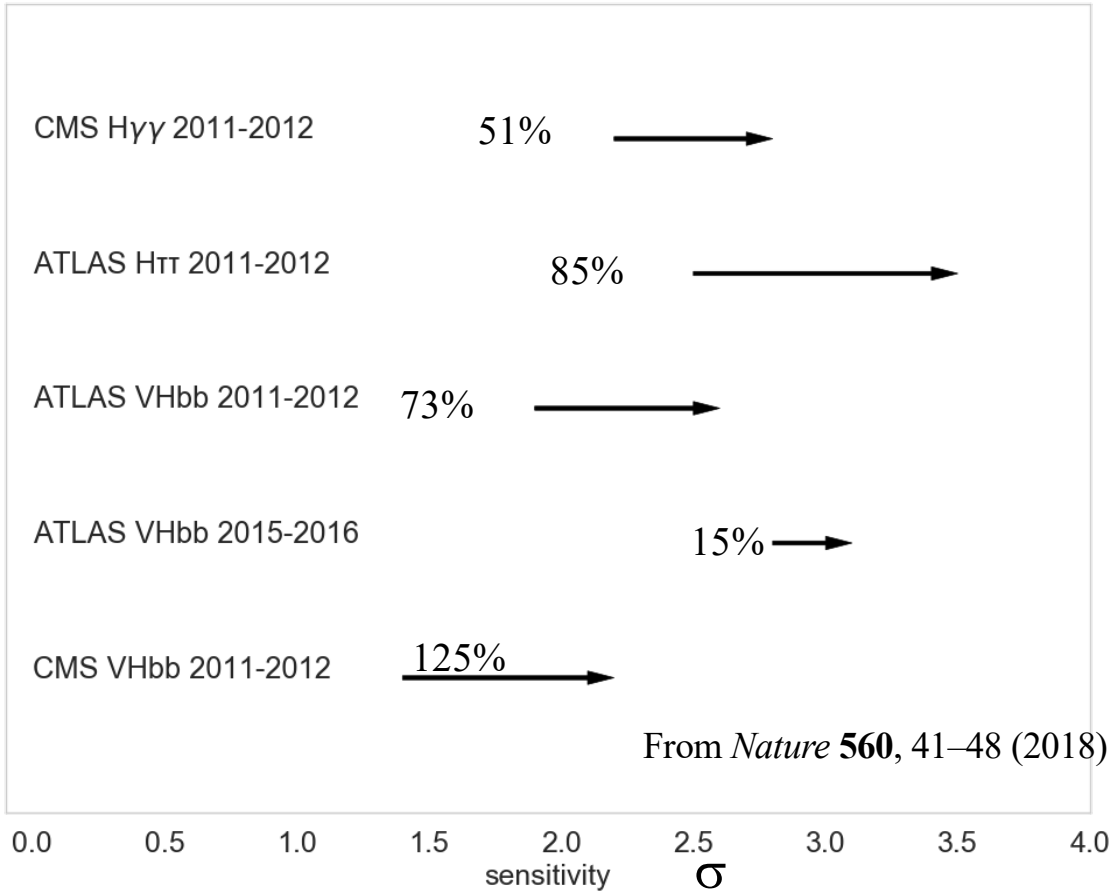
Higgs Boson at LHC with a classifier

JHEP 04, 117 (2015) 1501.04943

Boosted Decision Tree on a dozen features



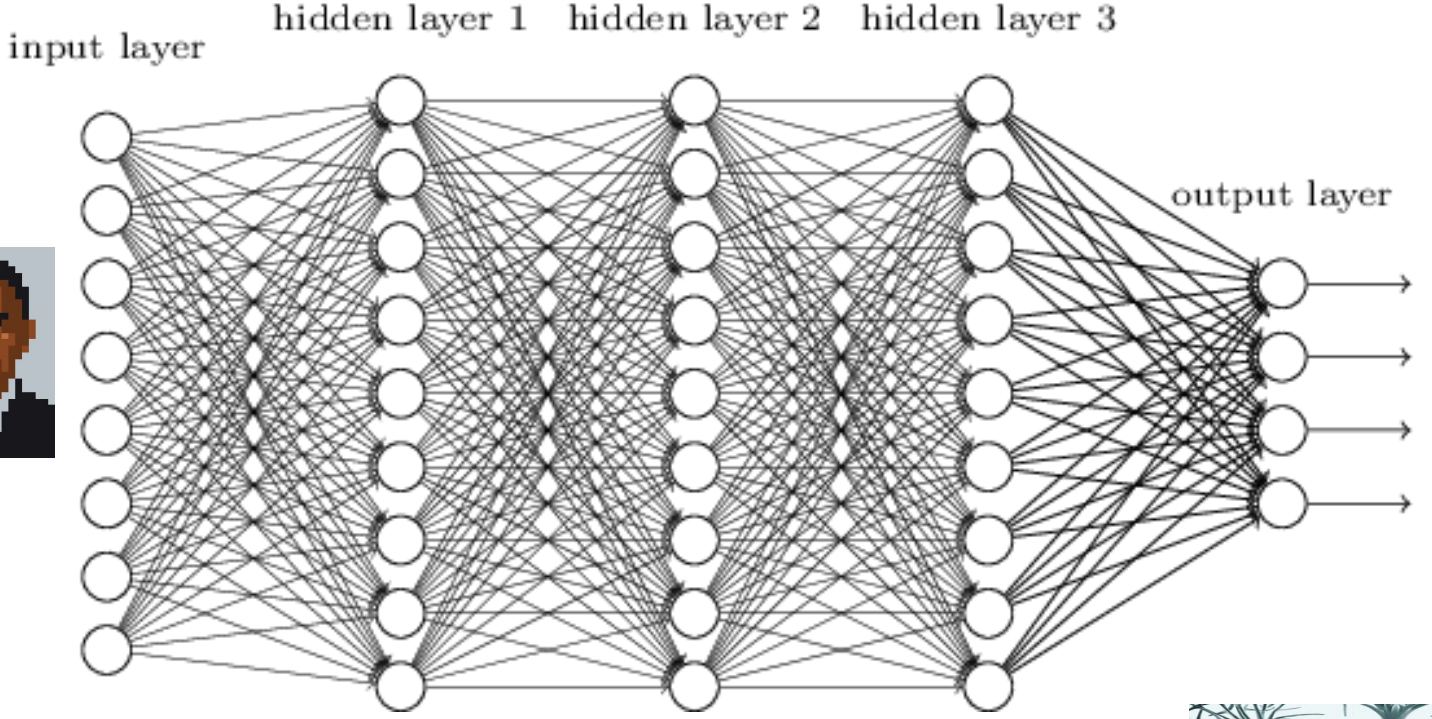
Impact of ML on Higgs Physics



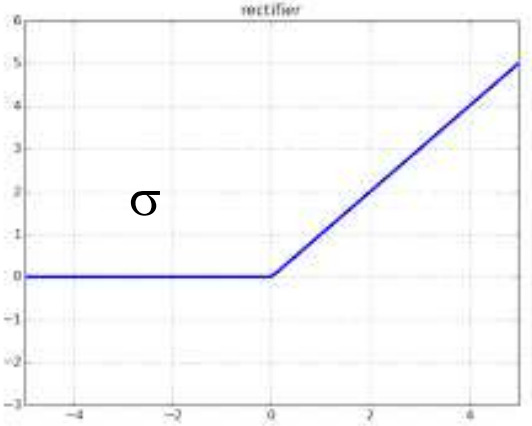
➔~ 50% gain on LHC data!

Neural Network sur des images

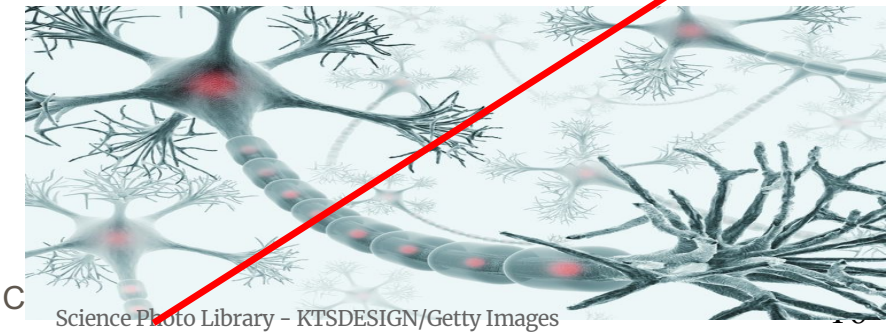
Neural Network



activation function (non linear!)



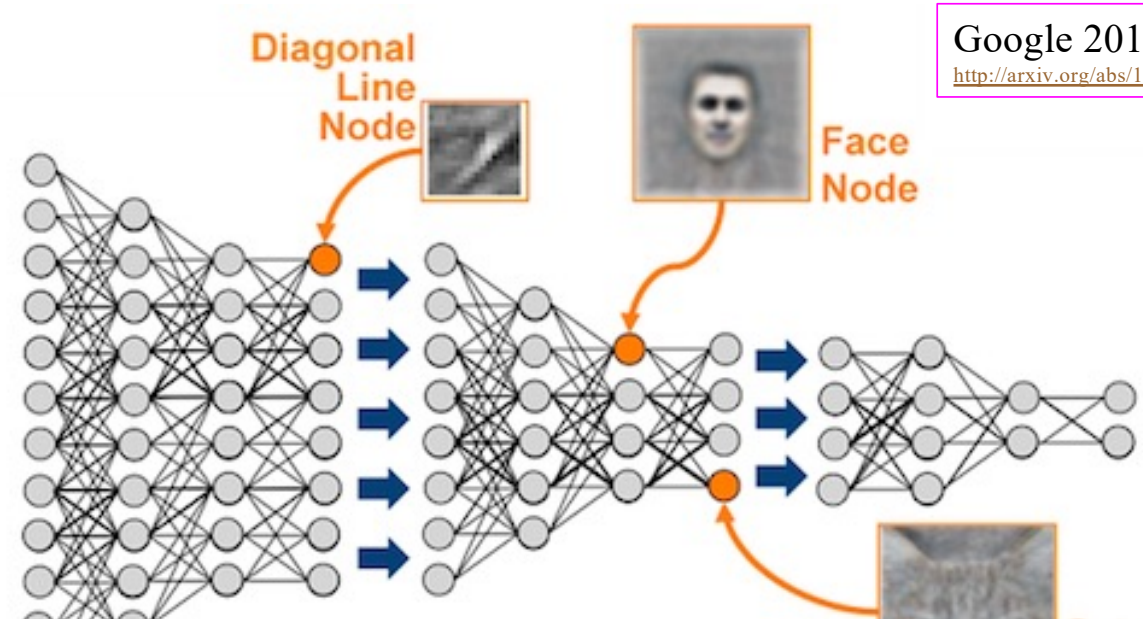
$$\text{Node}_j = \sigma\left(b_j + \sum w_{ij} x_i\right)$$

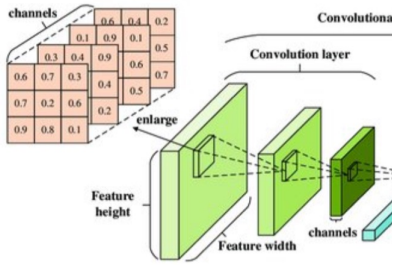


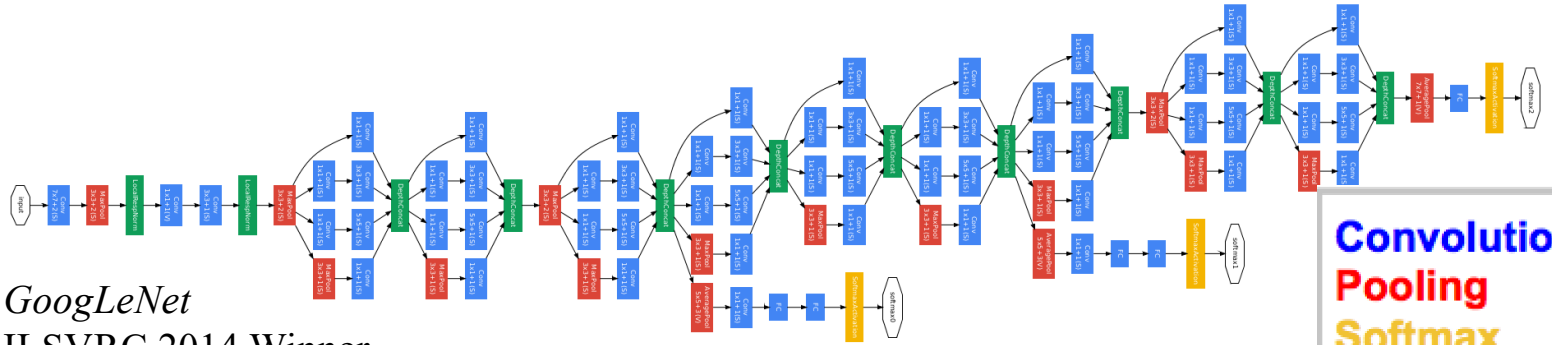
Science Photo Library - KTSDESIGN/Getty Images

Convolutional Neural Network

Google 2012
http://arxiv.org/abs/1112.6209



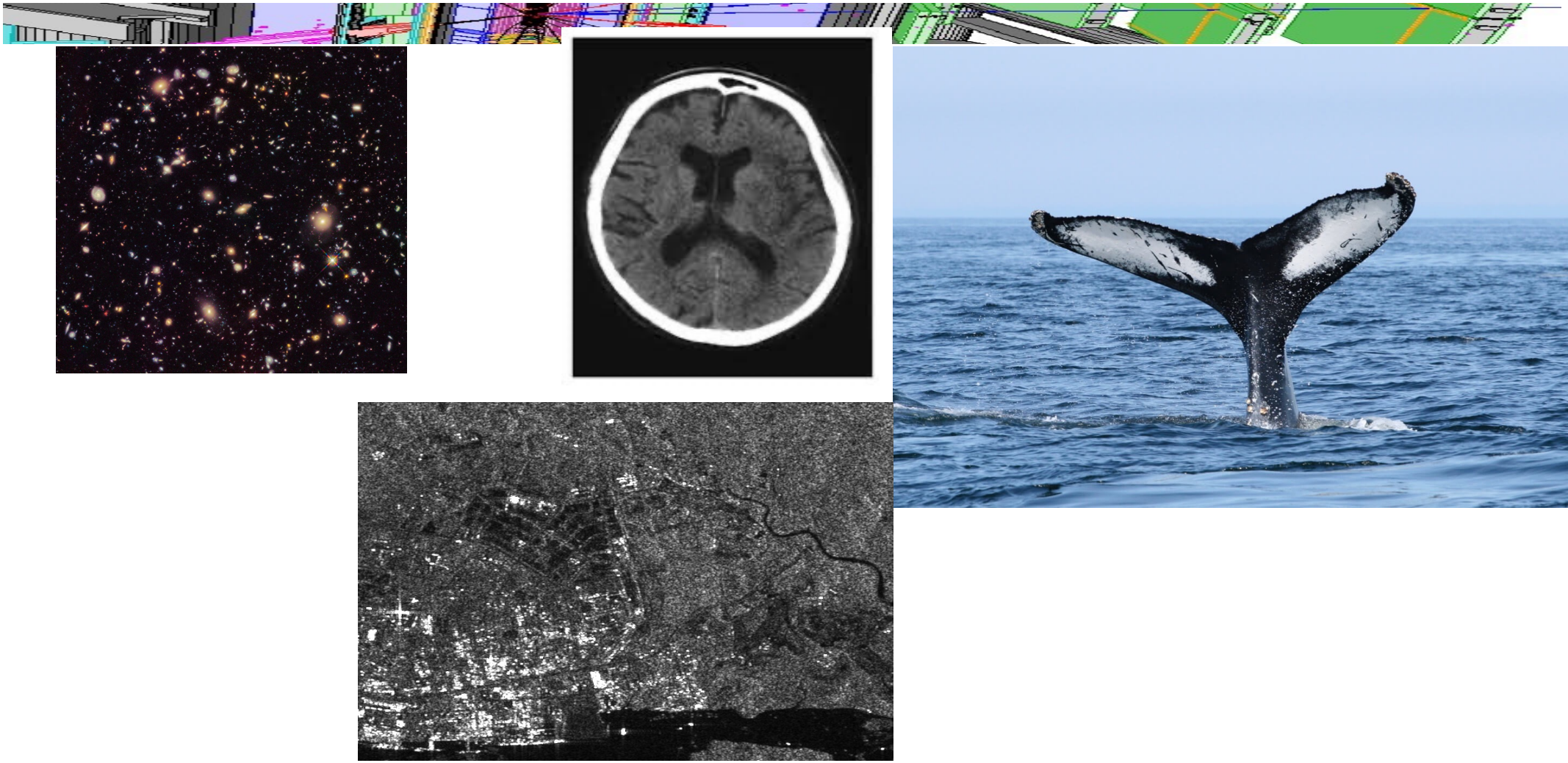




GoogLeNet
ILSVRC 2014 Winner
4M parameters

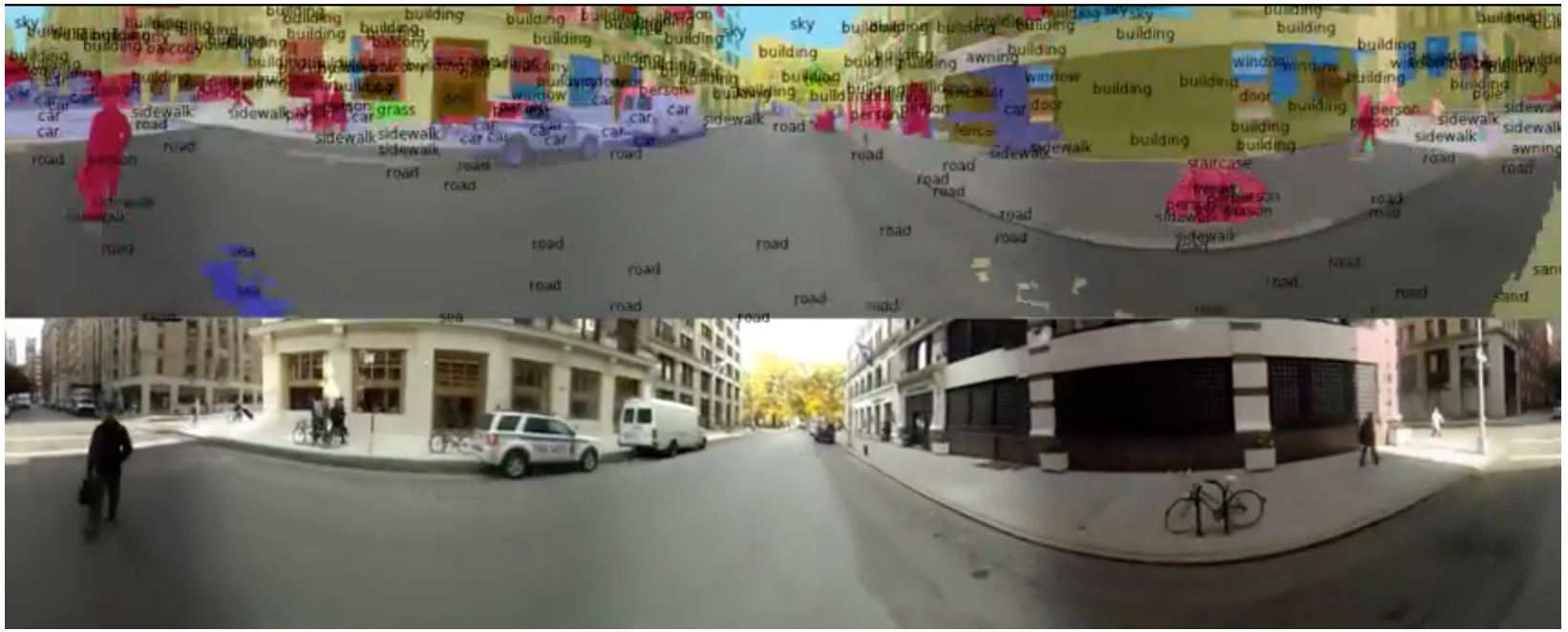
Convolution
Pooling
Softmax
Other

Many Applications



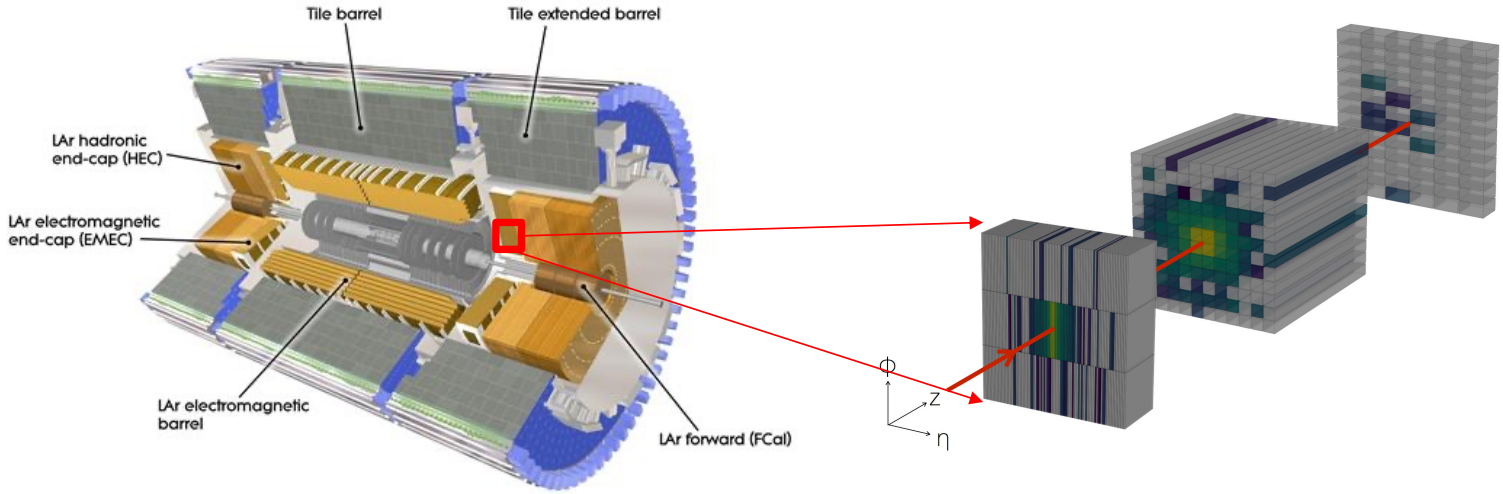
Scientific data are often not images

Typical Deep Learning application

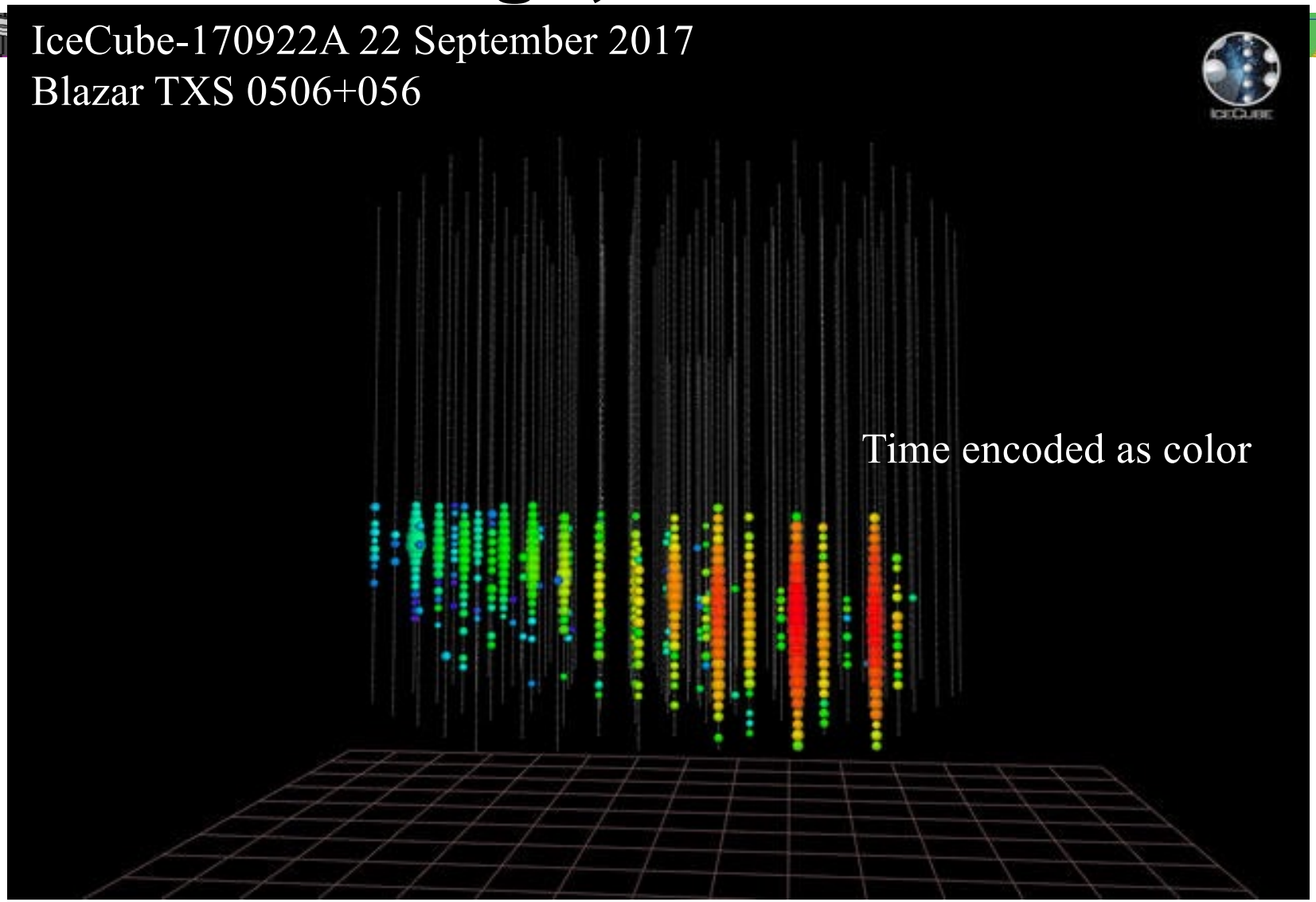


(supervised) learning needs labeled data➔accurate simulator are very useful!

Scientifica instruments are complex 3D objects

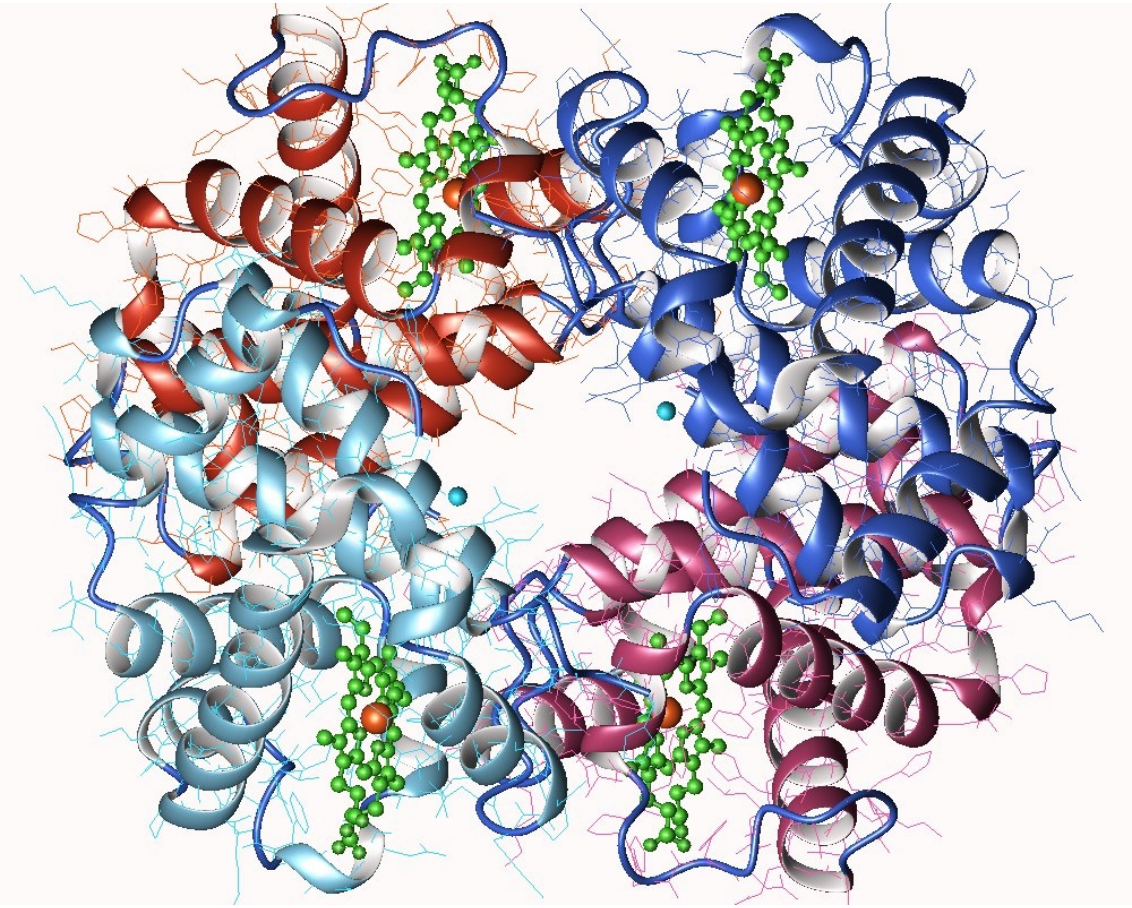


An image, not the data



An image, not the data

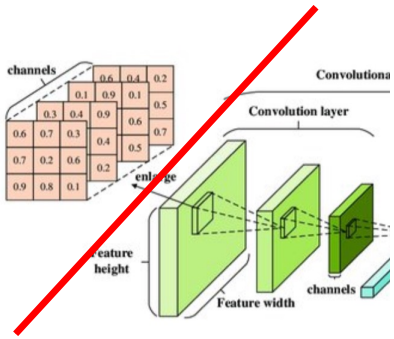
hemoglobin structure



Par Deposition authors: Fermi, G., Perutz, M.F.; visualization author: User:Astrojan — https://www.rcsb.org/structure/3hhb, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=48714519

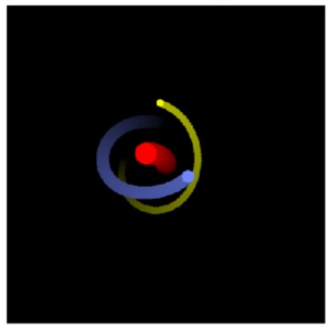
Specialised architecture

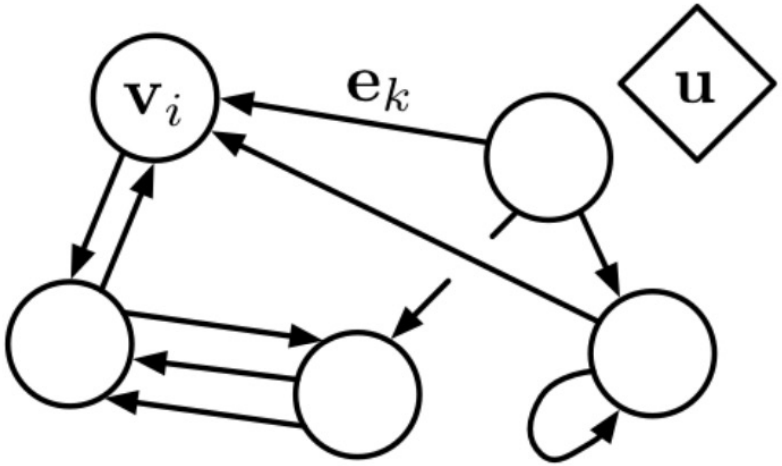
Graph Neural Network

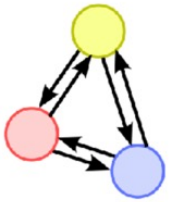


- Structure définie
  - v_i : noeud
  - e_k : arête
  - u : global

n-body







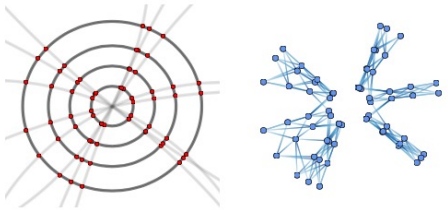
Nodes: bodies

Edges: gravitational forces

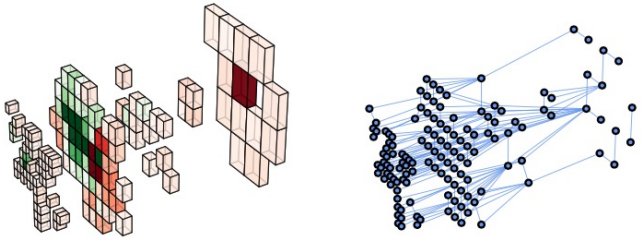
Global : potential energy

Graph on HEP data

from 2007.13681



(a)

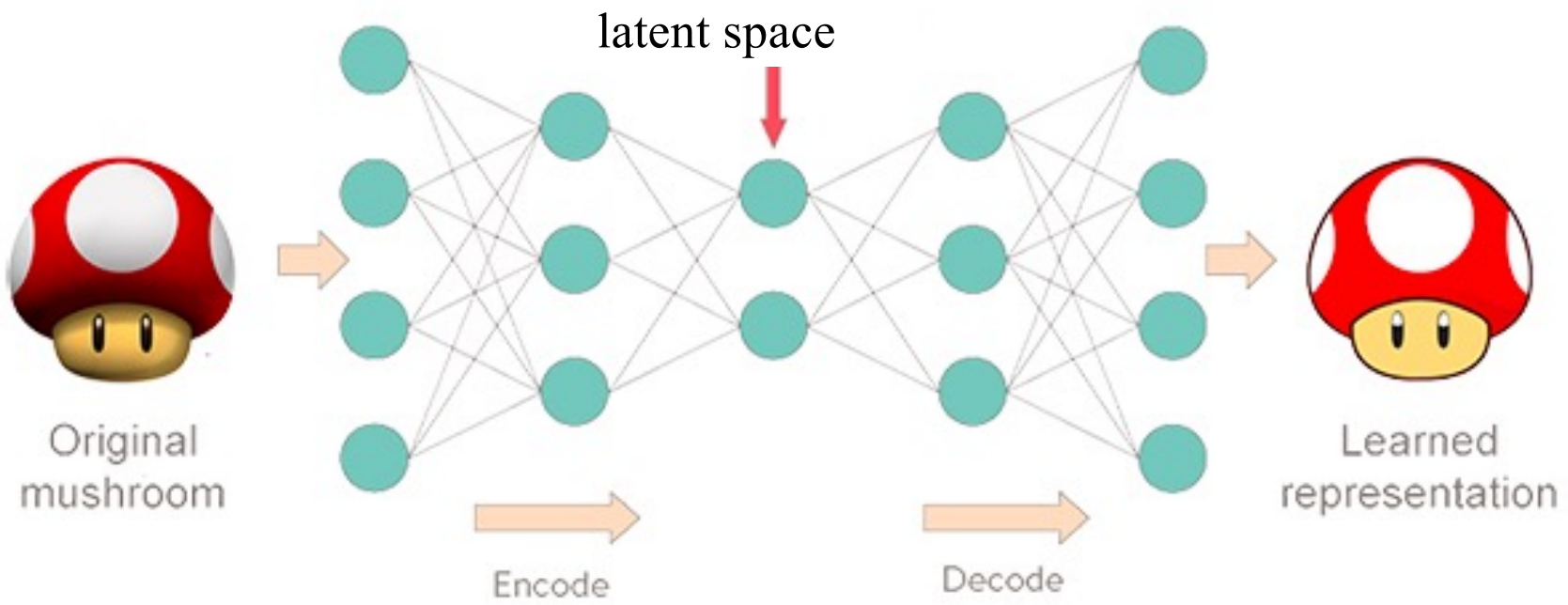


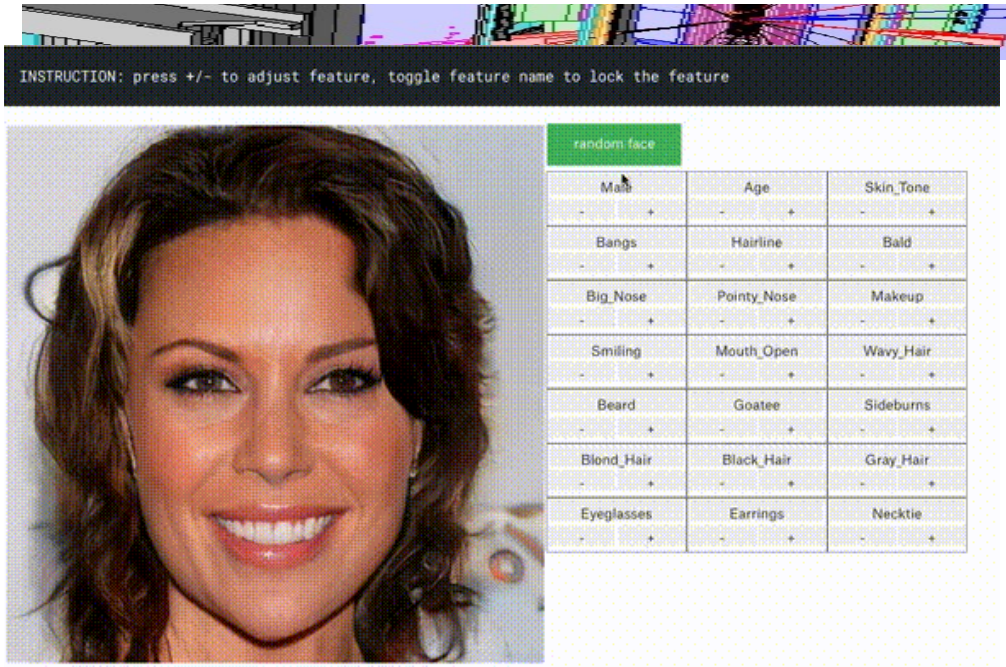
(b)

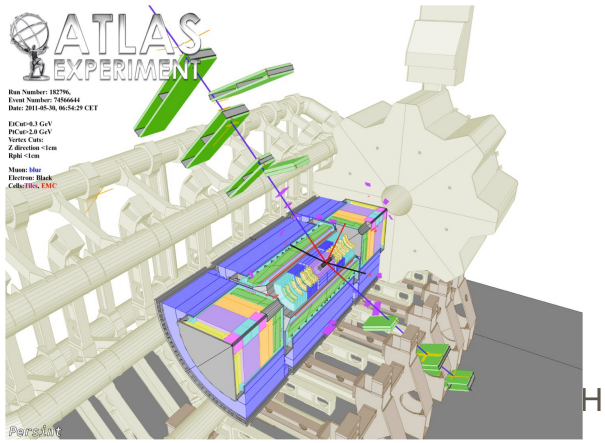
(c) (d)

Latent Variables

Auto encoders







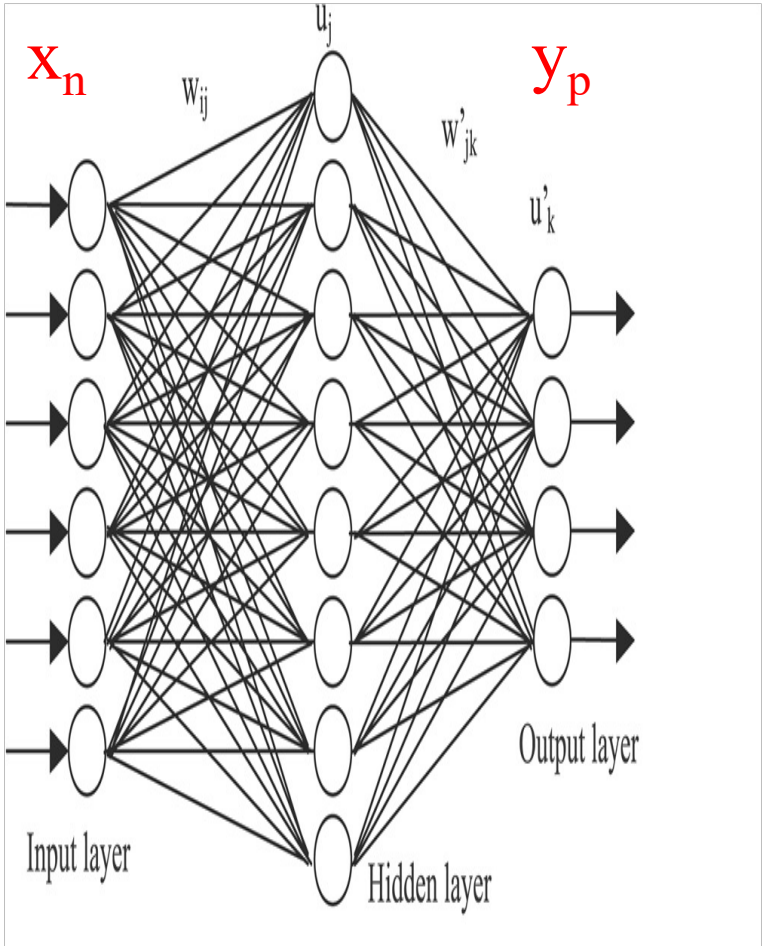
- latent variables==(few) abstract variables summarising differences between images
- Applications:
  - Classification/evaluation
  - Fast Generation
  - Explicability➔relation between latent variable latente and scientific concept
  - Inverse problem
  - Anomaly detection

Surrogate models

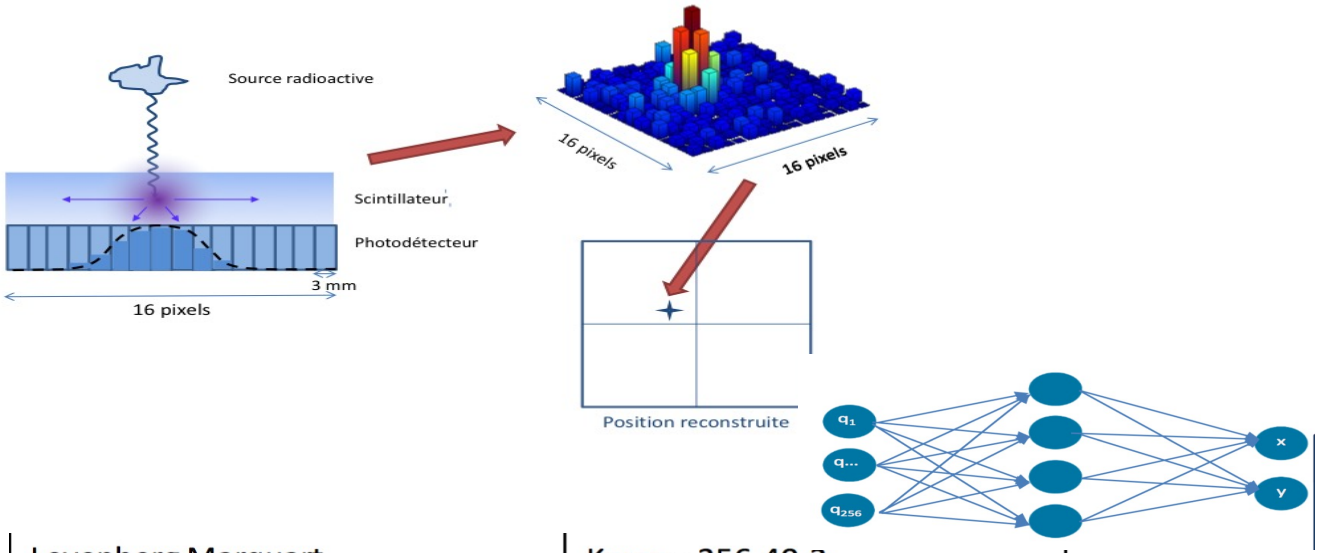
When AI does not better, but as well and much faster

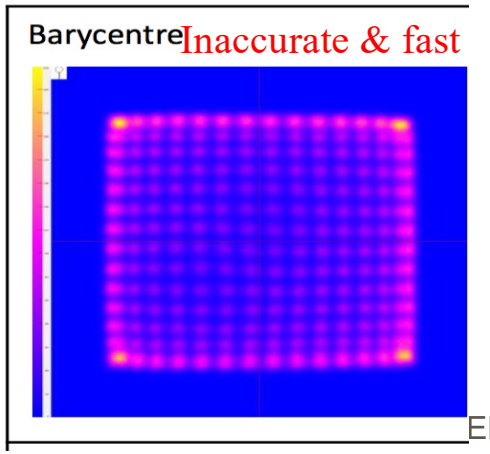
Universality theorem

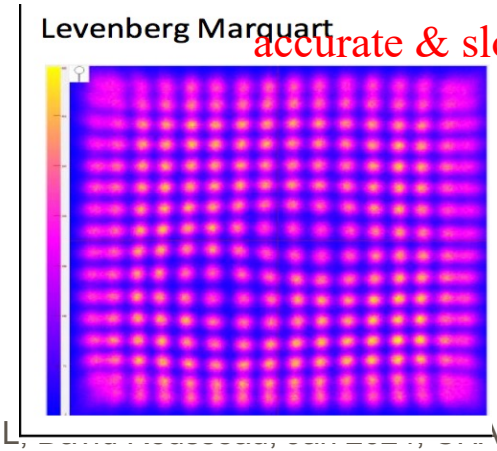
- Mathematical Theorem 1991
  https://en.wikipedia.org/wiki/Universal_approximation_theorem
- Any continuous function $R^n \rightarrow R^p$
- … can be approximate sufficiently well
- … with a NN with one hidden layer and enough node
- (does not say how to build this NN)
- ➔one can train a NN to emulate an existing simulator
- ➔speed gain x10… x100000
- (however precision…extrapolation…)

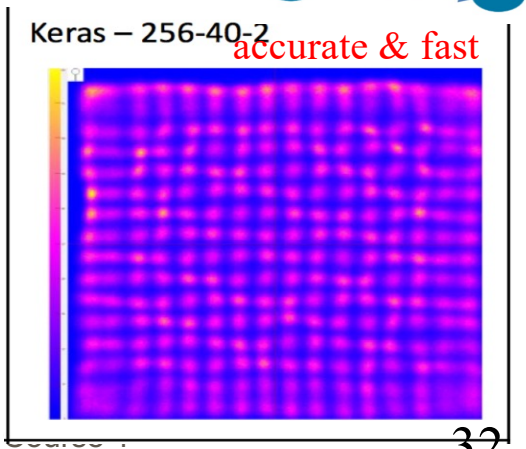


Exemple : β,γ camera for chirurgical

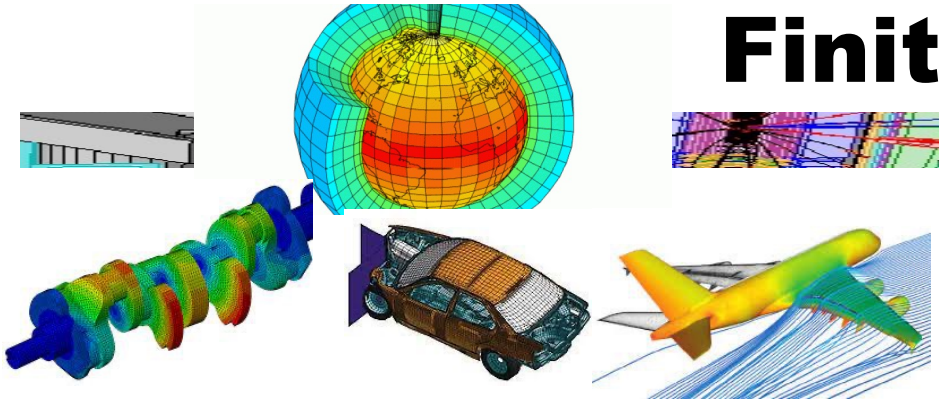


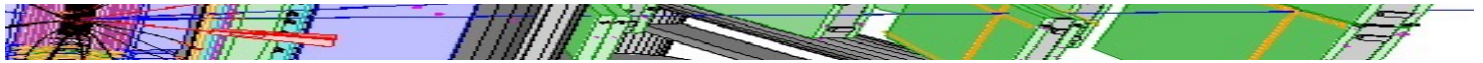




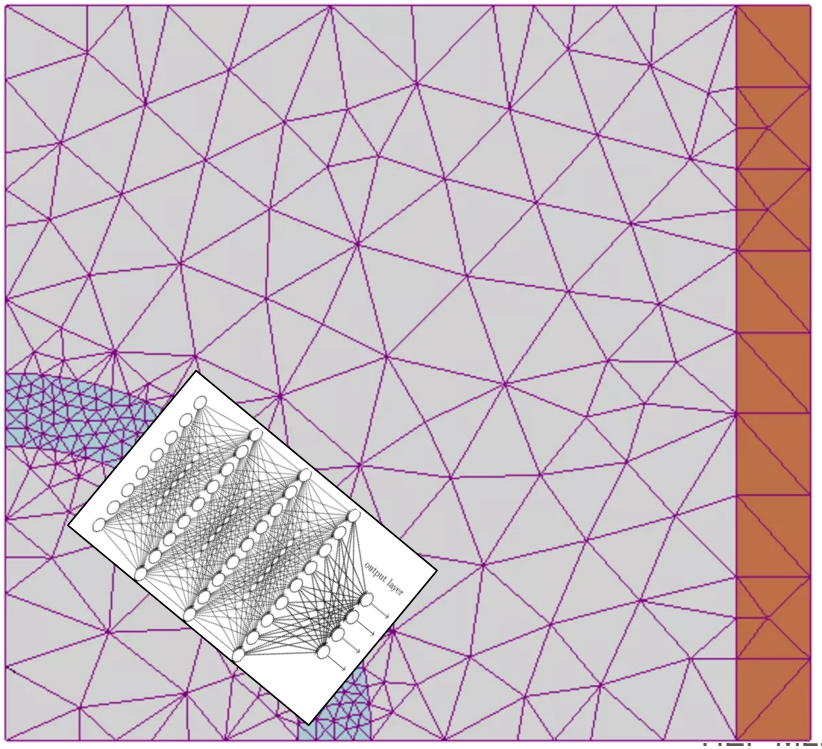


Finite Elements

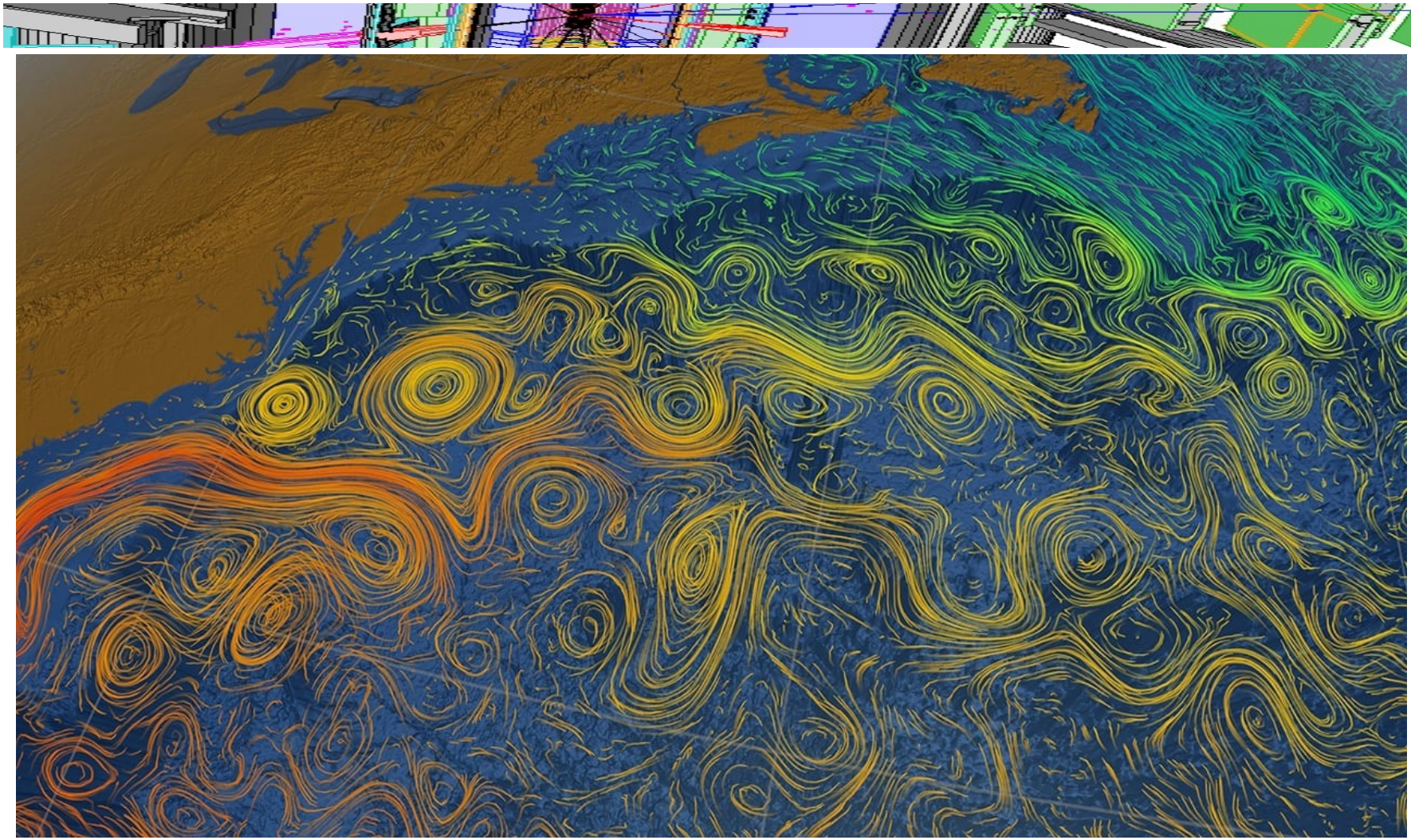




- Fundamental techniques
- Limited by mesh size
- maille / 2 ➔ resources x 8 (memory, speed)
- Instead of reducing mesh size, train an NN to emulate the behavior of a set of small meshes



Application oceanography



NASA

Modeling the unknown

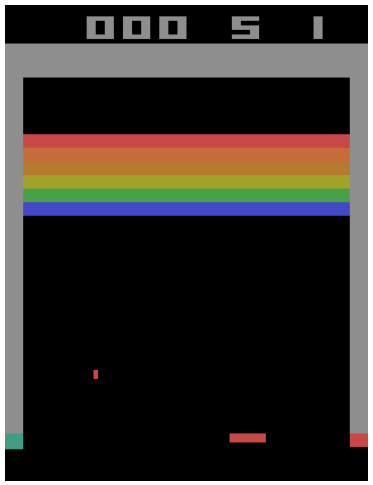
If a phenomenon cannot be modelled, emulate it with an NN

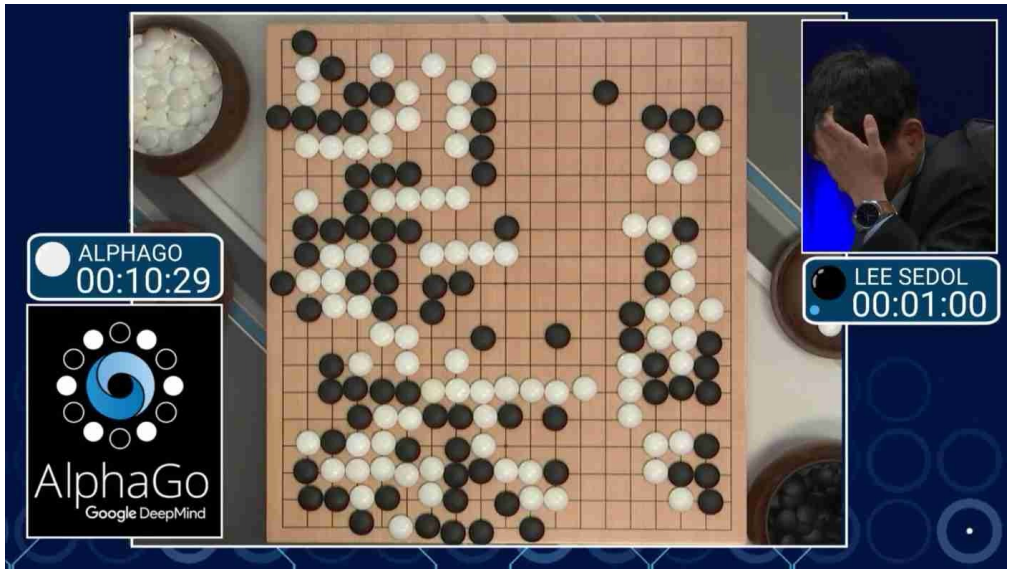
Ex: study of rogue wave Hokusai

Reinforcement Learning (RL)

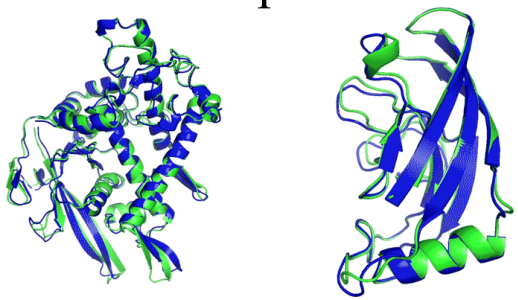
"It's not the destination, it's the journey"
Ralph Waldo Emerson

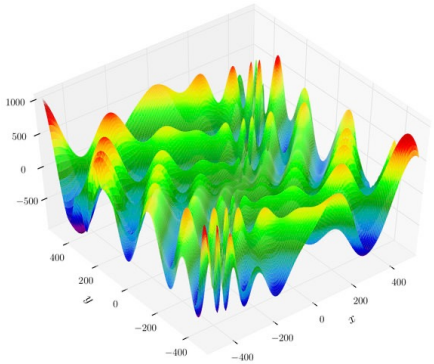
RL works better if





Alpha-Fold





- Classical methods fail
- Finite universe, but large combinatorial
- Clear objective, but unclear path
- Fast evaluation of trial