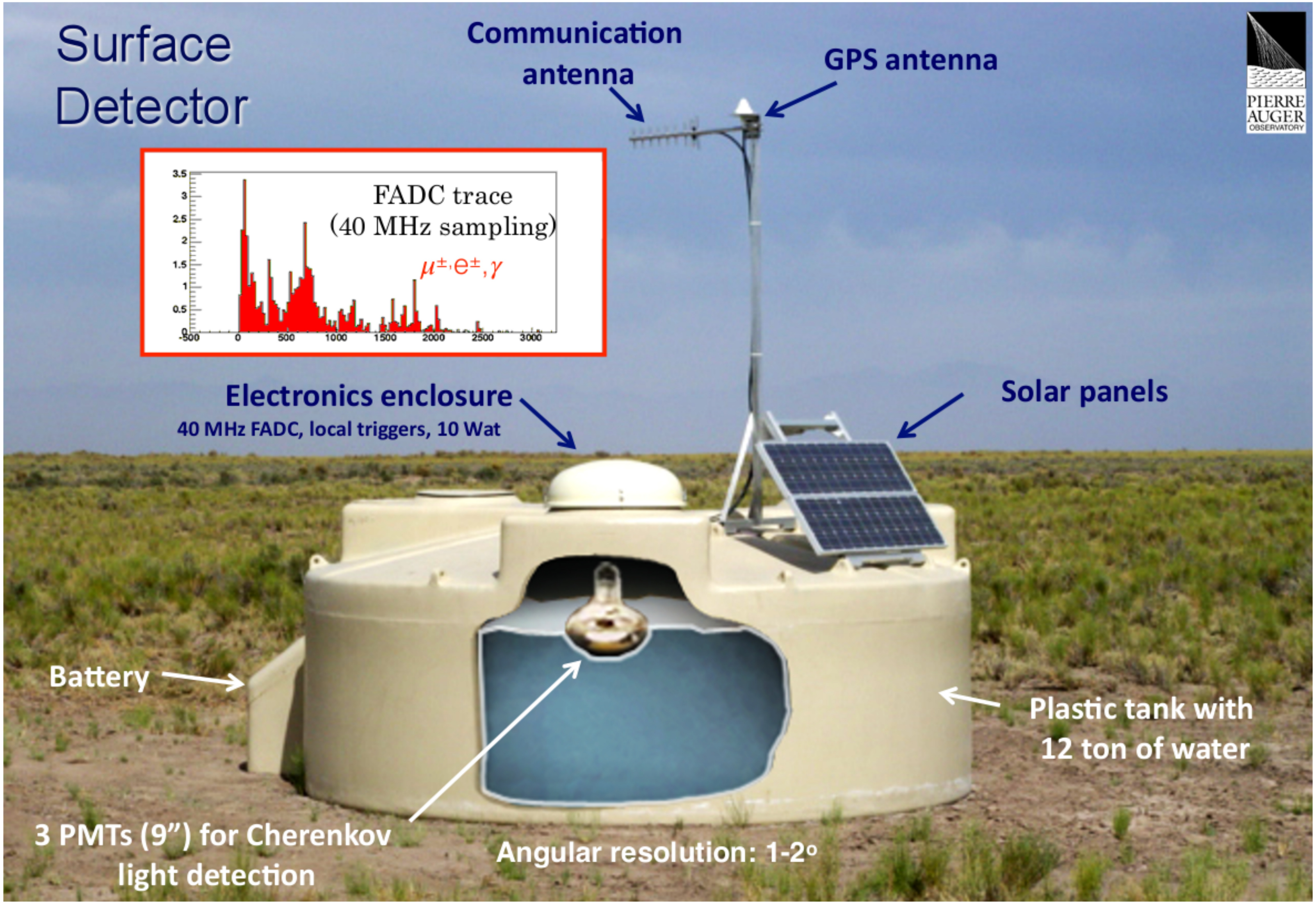
Photon classification in WCD using supervised and semi-supervised approaches

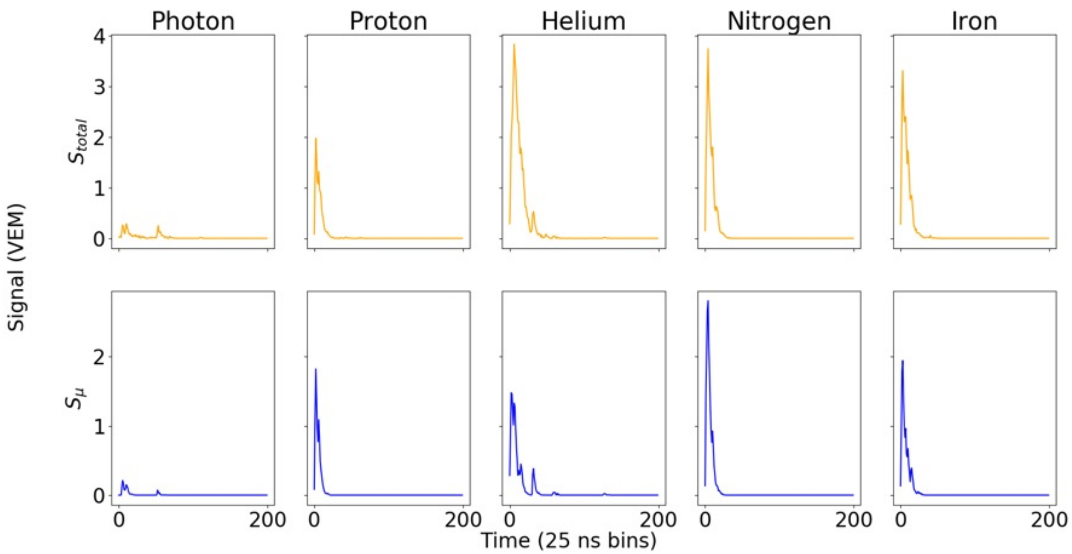
Alberto Guillén
University of Granada

Data and problem description



Data and problem description

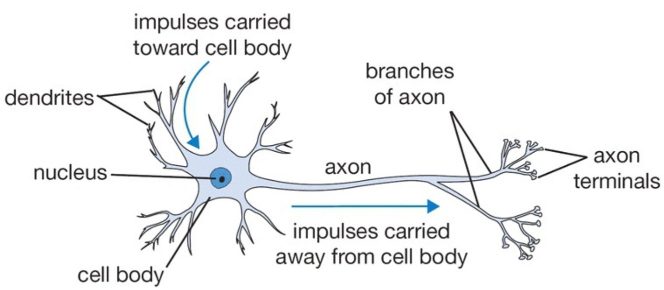
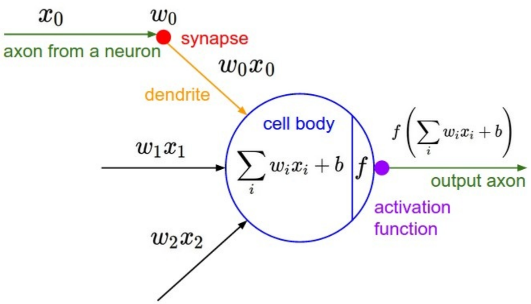
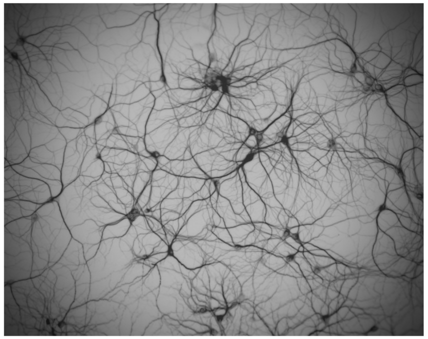
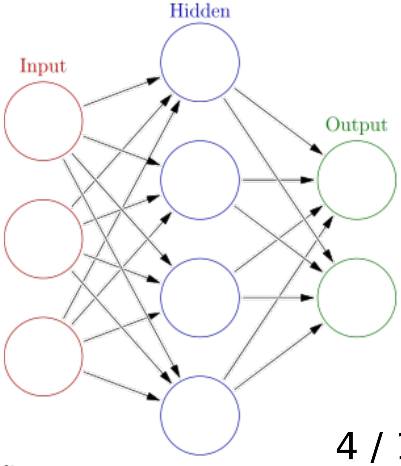
CORSIKA + QGSJET-II + Offline simulations from Pierre Auger Collaboration



Problems : **Muon number prediction**
Muon signal reconstruction
...
Binary classification: photon vs. Hadron
photon vs proton

ANNs, CNNs and DL...

- Simplified approximations of biological neural networks
-

	Natural	Artificial
Synapsis		
Neural Network		

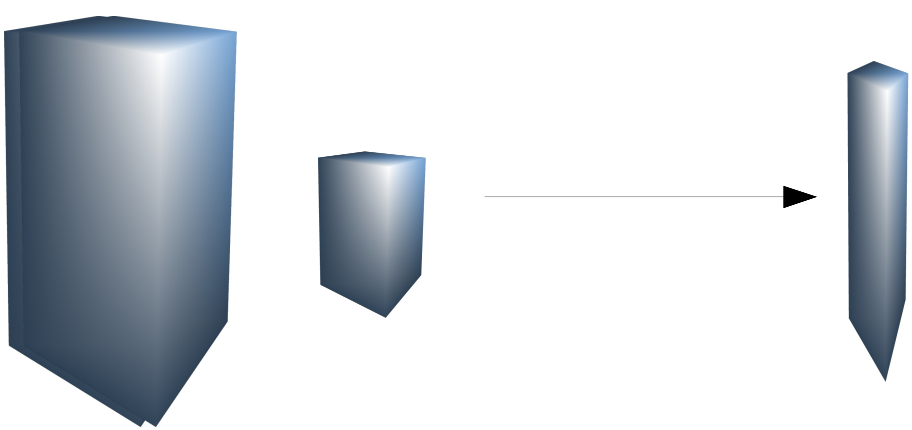
$g(x;\Theta)$

$\sum g_i(x;\Theta) W_i$

Images:
https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSbb2Van-e2T24h3Z44c-HfUr4PXu-LcCNs3Gg2OVdT3_aY1dR9ng
http://cs231n.github.io/neural-networks-1/
https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQlG8BgVOI8a7mrwnj9X0F8p1Q8Nj2umPnWy5-kEaskuXfEuWdn

Convolution (in DL)

- Given a 3D tensor (height,width,depth), a filter (receptive field) operation (<height, <width,=depth) is applied
  - Proyected 2D tensor (activation map) is obtained conserving *presummed* local relationships



The more filters, the more convolutions can be “combined” afterwards to stack more conv. layers

Problem approximation 1: only using the trace

- Input: 3 WCD traces per event (with signals S_{total}, S_M, S_{em})
  - Some approaches:
    - classify isolatedly each station (= One trace)
    - Use *three* traces as input
    - Using S_{total} and S_M (AugerPrime will provide both...)

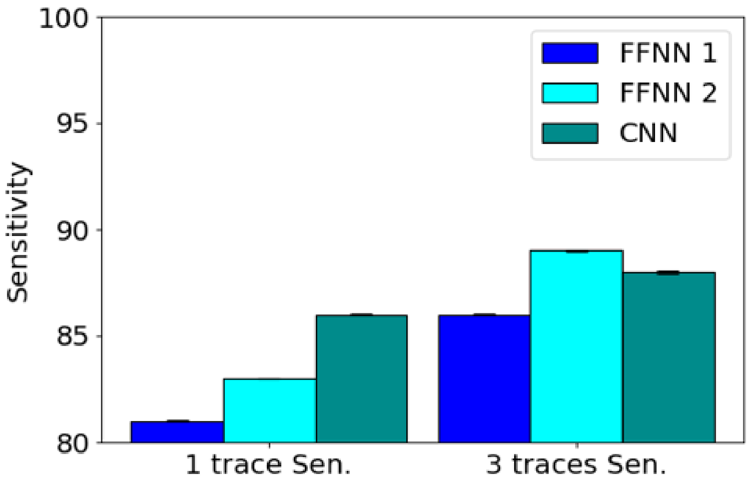
5 types of particles → Imbalanced classification → 1st approach: undersampling hadrons

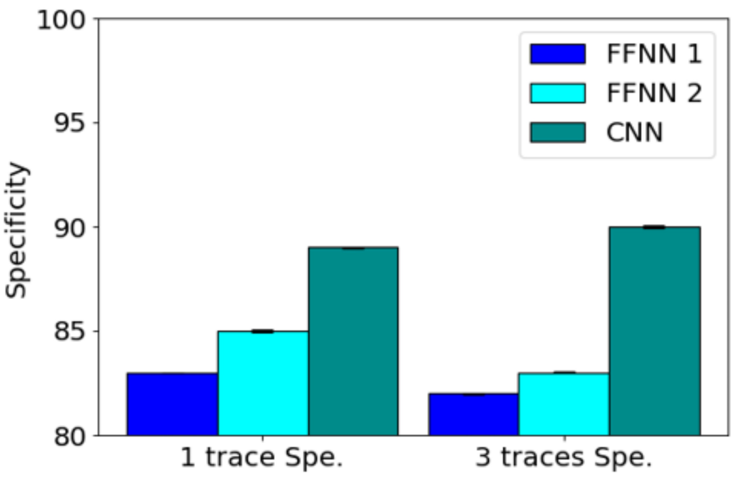
	Photon	Hadrons	Total
Using 1 trace/event	39195	40000	79195
Using 3 traces/event	3312	3600	6912

Results using ANNs

Three ANNs architectures evaluated using the total signal and muonic signal

One trace			Three traces		
S_{total}	**Test Acc.%**	**Val. Acc.%**	S_{total}	**Test Acc.%**	**Val. Acc.%**
FFNN1	82.12(0.65)	82.11(0.27)	FFNN1	84.03(1.40)	86.20(0.68)
FFNN2	84.14(0.43)	84.18(0.17)	FFNN2	85.79(0.80)	88.28(0.58)%
CNN	87.23(0.22)	88.74(0.14)	CNN	88.48(0.02)	90.95(0.58)
S_{μ}	**Test Acc.%**	**Val. Acc.%**	S_{μ}	**Test Acc.%**	**Val. Acc.%**
FFNN1	78.15(0.074)	78.55(0.37)	FFNN1	87.08(0.41)	86.98(0.65)
FFNN12	77.92(0.074)	78.61(0.37)	FFNN2	86.27(0.41)	87.86(0.95)
CNN	79.05(0.12)	79.05(0.34)	CNN	86.07(0.34)	88.02(0.60)





First preliminary conclusions...

- Photon identification seems feasible just using the (simulated) trace
- Spatial and temporal information intra and intertrace is useful (better use three traces) for CNNs
- Using total signal is better than considering only muonic → S_{em} might be a good input...

Conclusions

- Given a set of observations, unsupervised learning seems not possible (with clustering)
  - Semi-supervised it does a good job gaining interpretability...
- For ultra-accurate results, state of the art models excel...

… as long as simulations are correct.

- Useful? For sure! if not for Physics, for design stage → models can quantify changes in design decissions

Thanks to

- LIP
- Pierre Auger Collaboration

...and all of you for your attention.