Neural Network for Position Reconstruction

and other problems

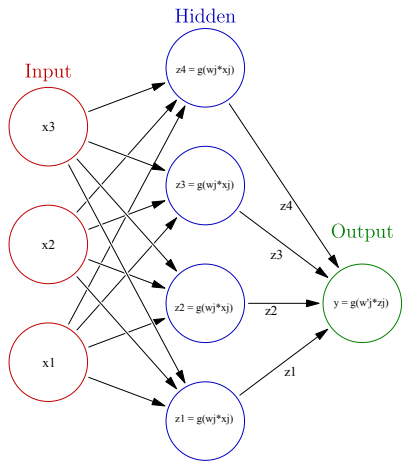
Miguel Cárdenas-Montes, Roberto Santorelli,
CIEMAT Team

CIEMAT

Madrid, 22 September 2023

Neural Networks and type of problems

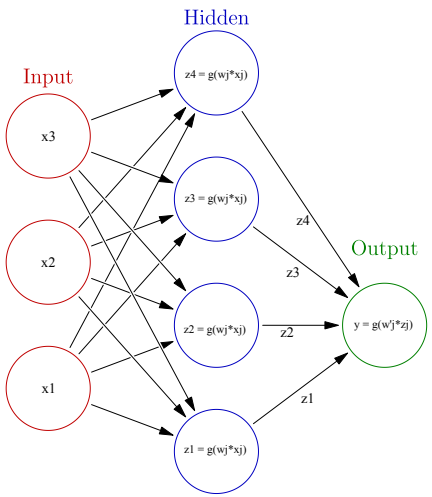
- A Neural Networks (NN) is able to **learn patterns** on labeled data (Montecarlo) and *guess on unlabelled and previously unseen data* (supervised learning).
- Position reconstruction is a regression problem in Machine Learning (ML) in which **continuous labels** (coordinates) are predicted.
- A second category of problems involves classification, in which **categorical labels** are predicted: such as background signal classification and particle identification.



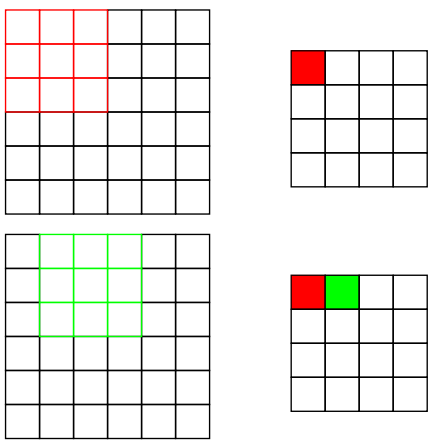
- In non-supervised learning unlabelled data are **clustered** by similarity.

Architectures, Discussion, pro-cons

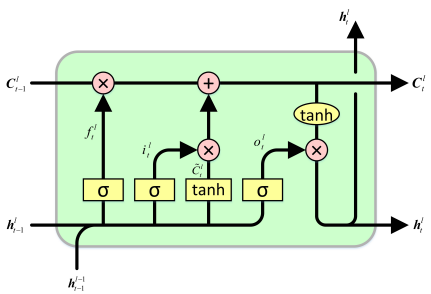
- NN Architectures: trainable parameters, pattern complexity to learn, time processing (trigger), etc.
- Correlation exploitation: temporal and/or spatial. Richer/more complex signals.



(a) MLP



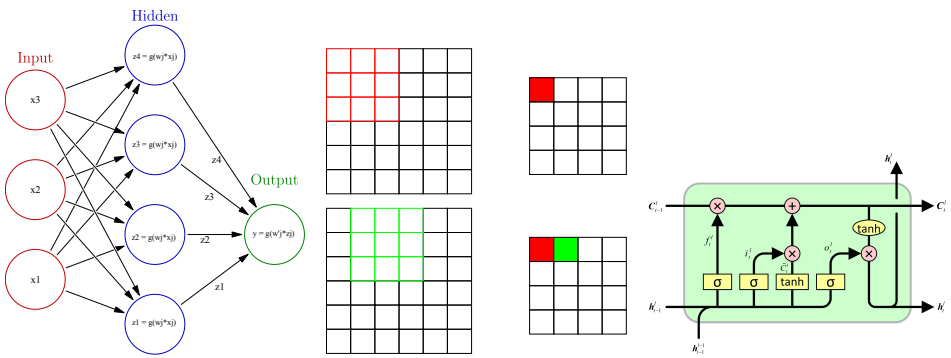
(b) CNN



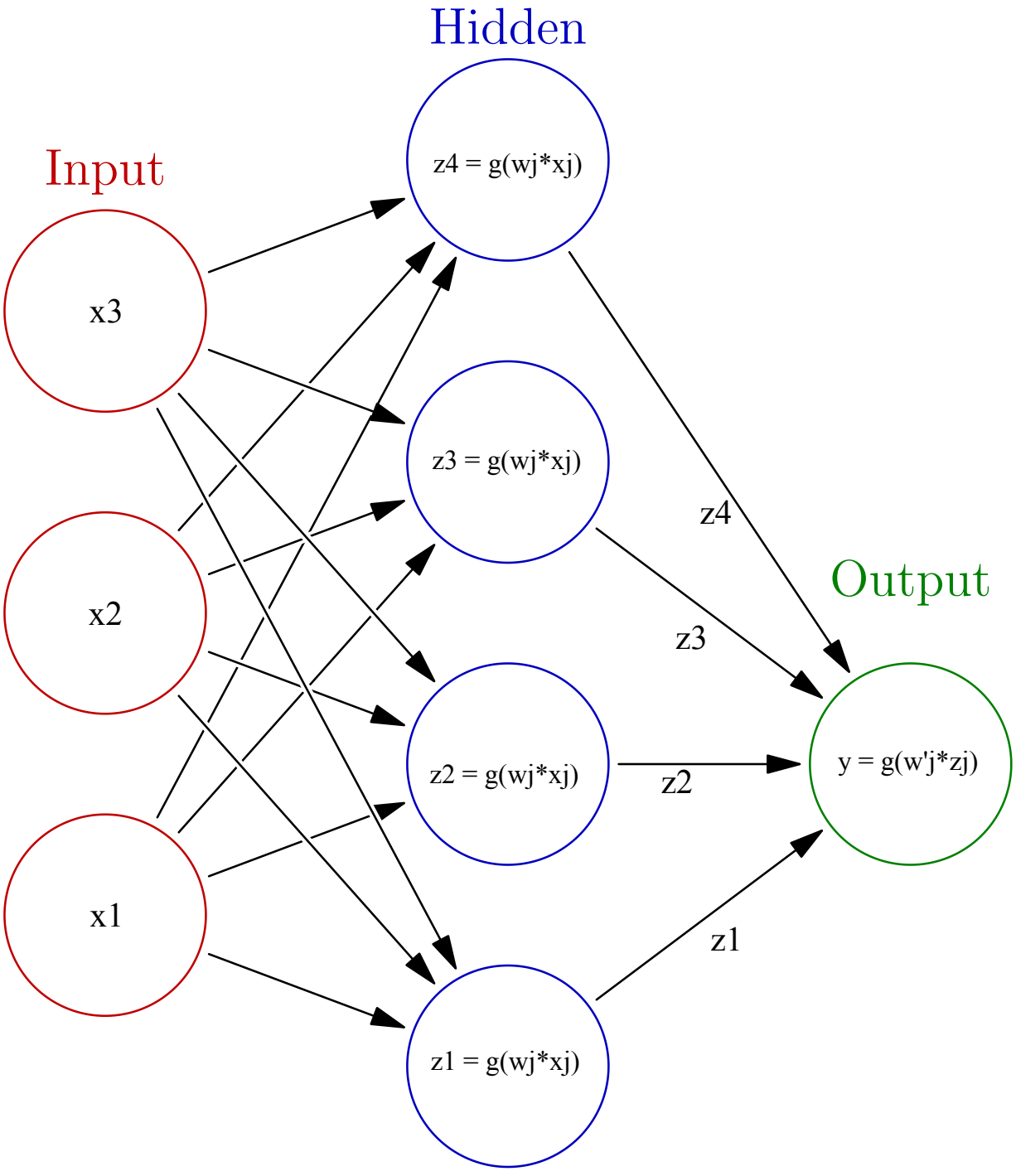
(c) RNN

DL can help on ...

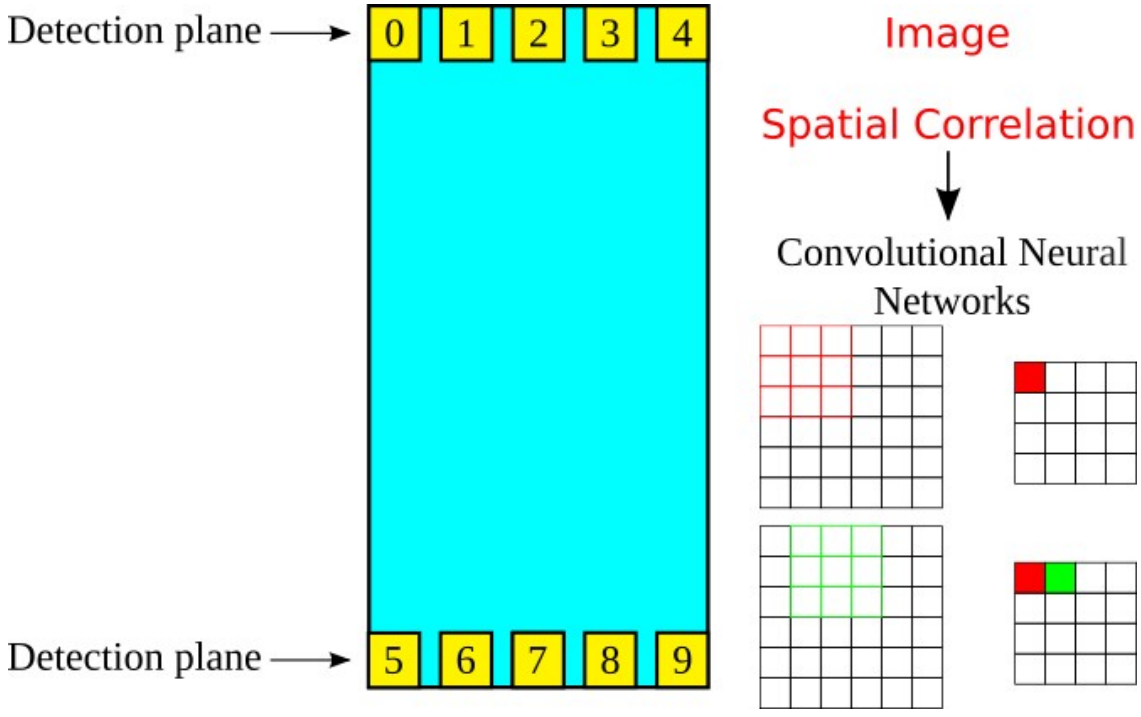
- As far as liquid noble gases TPCs become larger and larger, drift time takes longer, probability of pile-up increases.
- Therefore, to address particle identification (background rejection), accurate position reconstruction and multiplicity labelling NN/DL become mandatory.
- Detector are complex machines. Fail and close to fail regimes can be predicted (DL).



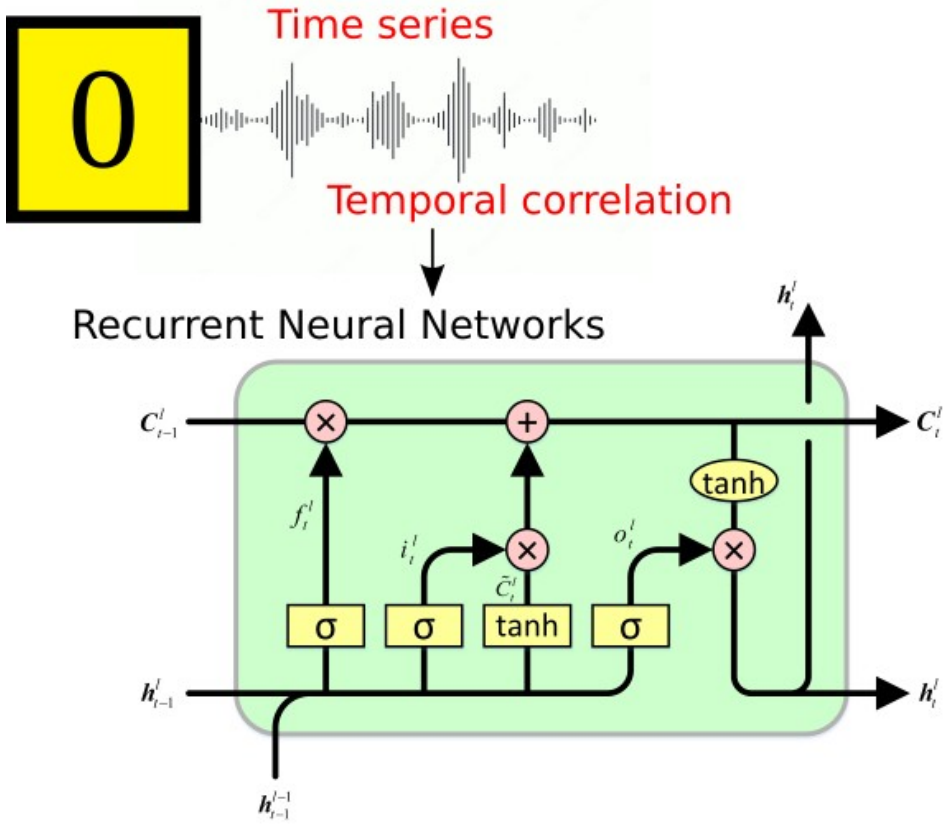
(a) MLP (b) CNN (c) RNN



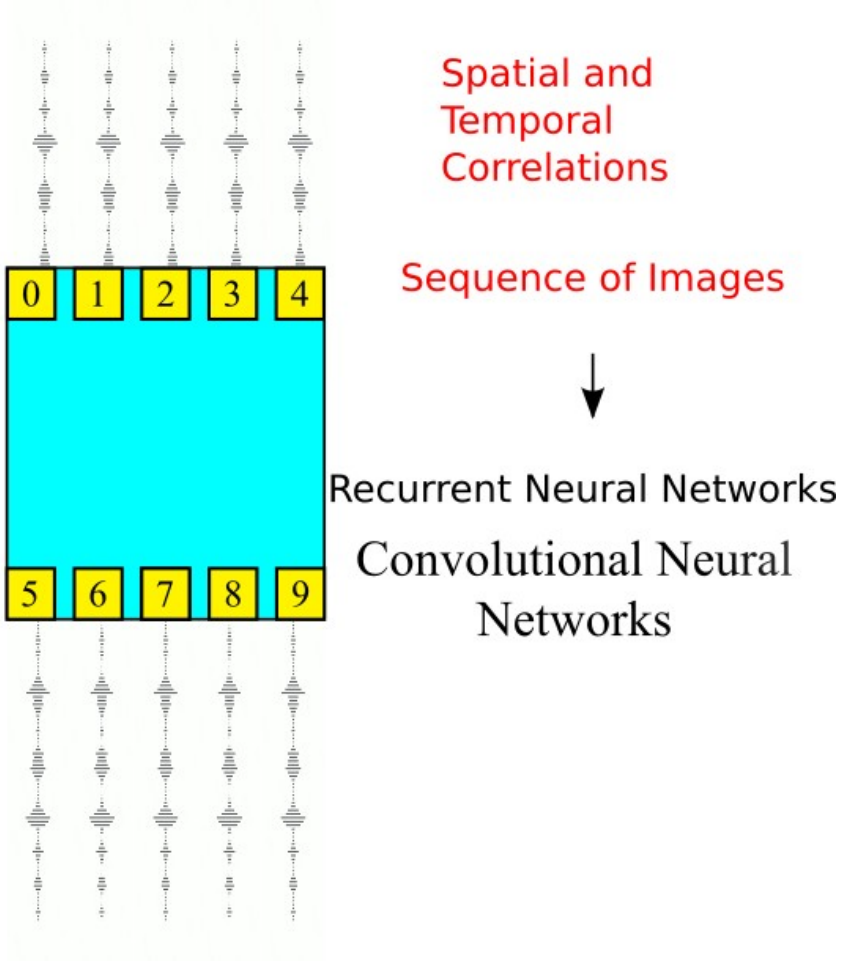
- Multilayer Perceptron (MLP) is not able to exploit correlations: neurons are a non-linear function wrapping the dot-product of variables and weights (trainable).



- Convolutional Neural Networks (CNN) are able to exploit spatial (topological) correlation on signal, both in 2D images, and 3D (true or fake) colour.
- CNN can also handle layers of spatial images (2D) arrange in a 3D tensor of arbitrary third dimension (multiple time steps of a 2D image, sequence of images).



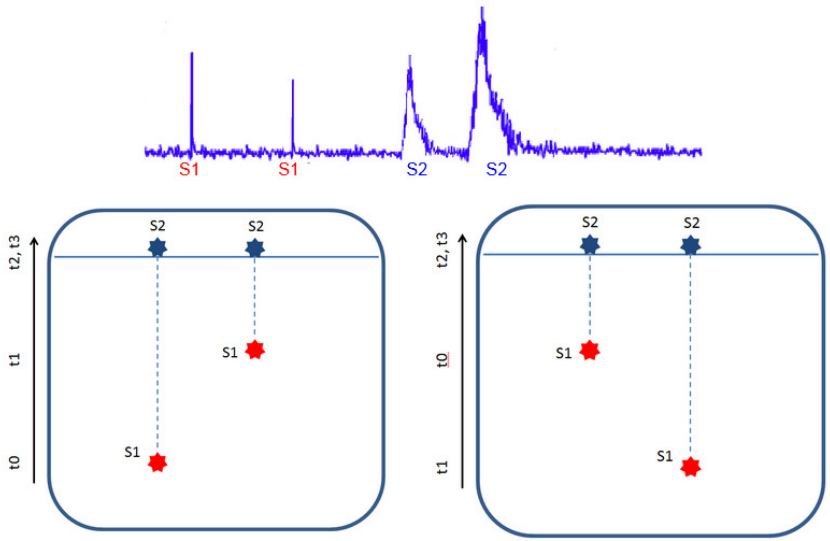
- Time series (pulse shape) can be used as discriminator by Recurrent Neural Networks (RNN): particle identification, but also as anomaly detection (close to fail prediction for subsystem at least).
- More complex architecture than CNN, thus more intensive computation and it takes longer.



- Much more complex architecture than CNN and RNN individually, thus it takes much longer.
- Large amount of data stored per event (computational resources).
- Data layout becomes more and more relevant for final performance.

Position reconstruction and pile-up

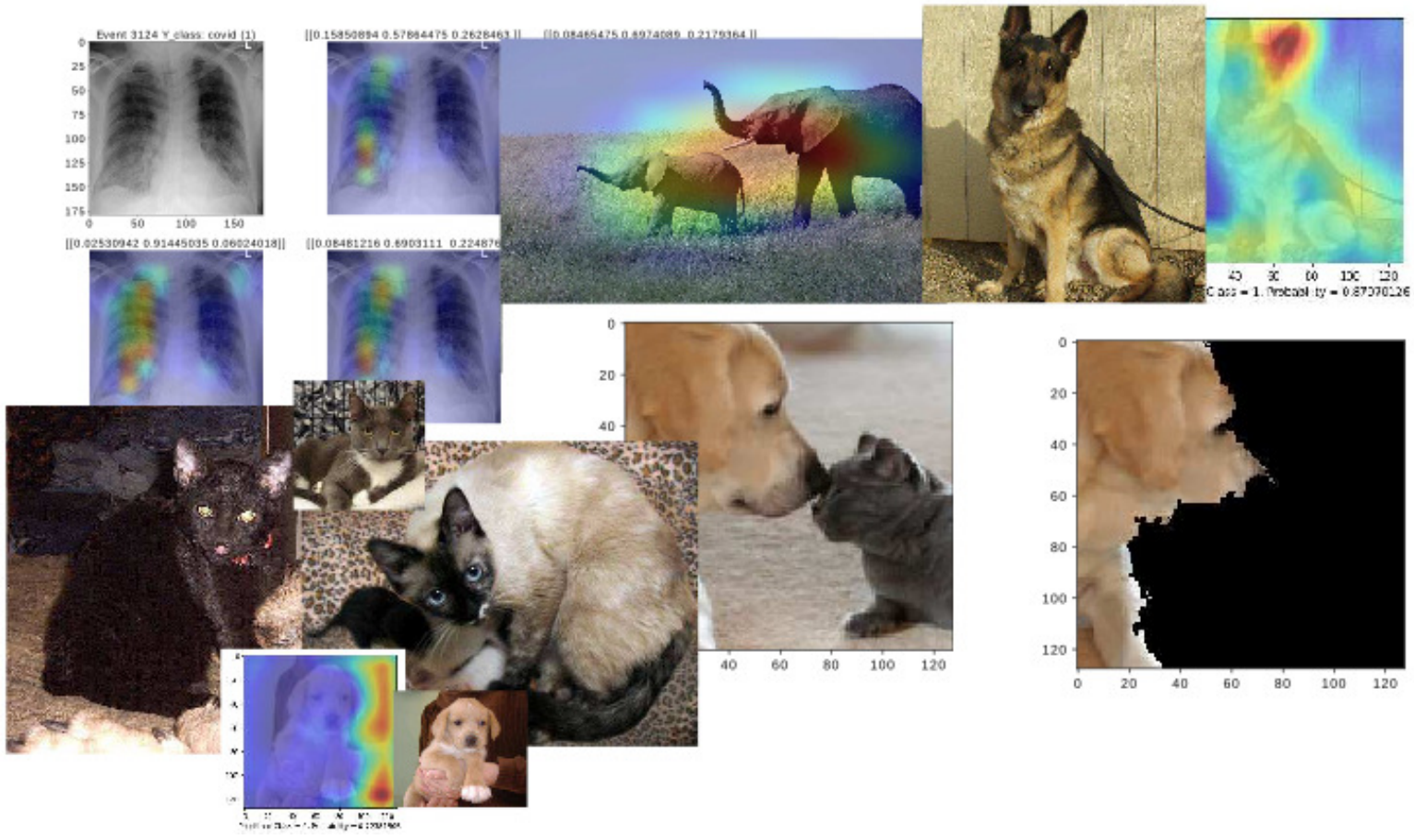
- In case we are able to reconstruct the z (possibly x,y) coordinate from the S1 light pattern, we can unambiguously associate S1 and S2 if the events are spatially separated.
- In this case, we have a certain time window in which the S2 corresponding to a specific S1 is expected.
- This could considerably relax the maximum event rate we can handle and the requierements on the event rate/material contamination.



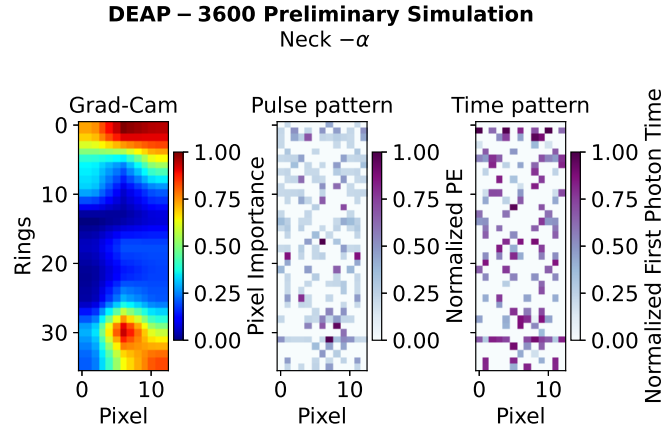
- Accurate S1-based position reconstruction could mitigate the problem associated to pile-up.

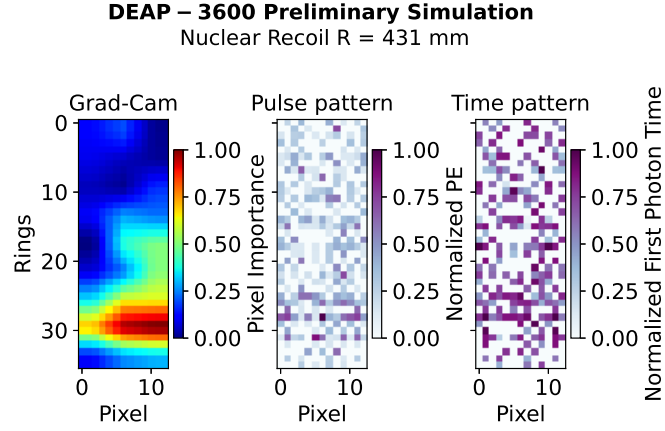
eXplainable Artificial Intelligence (XAI)

NN are not black boxes, unboxing



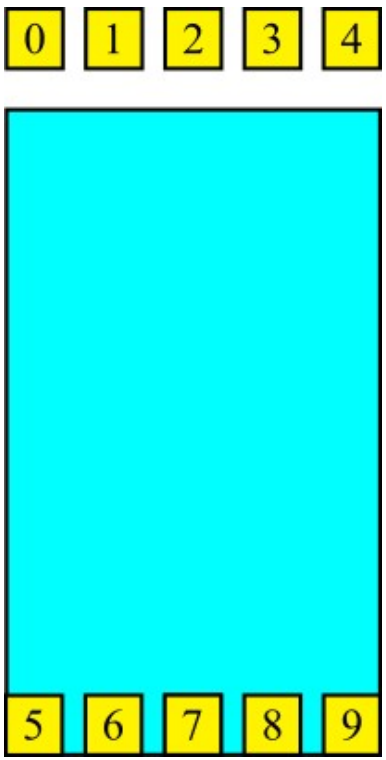
- XAI algorithm depends on input shape (1D, 2D) and NN architecture.
- It allows discovering salient features and bias in the dataset.

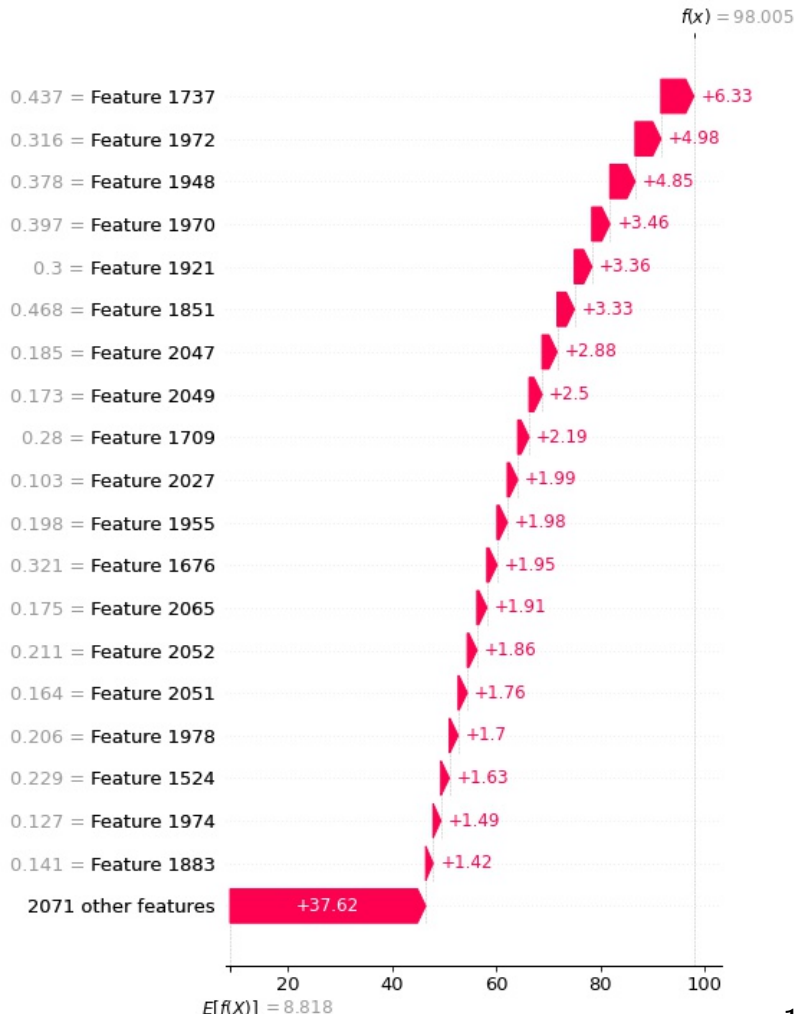




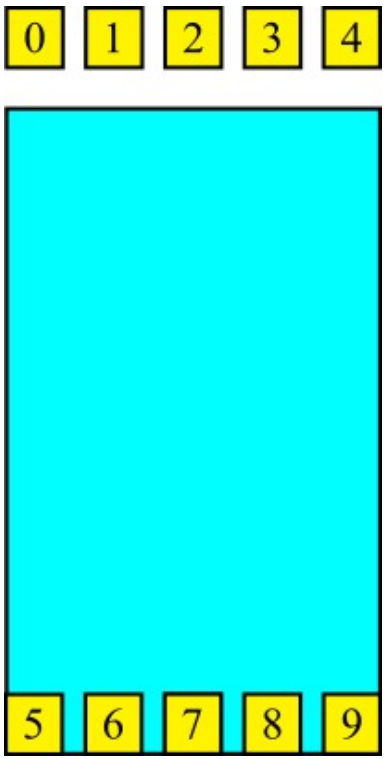
- Improvements of the performance by deeper architecture, larger dataset (more statistics), but also data representation (and the architecture).

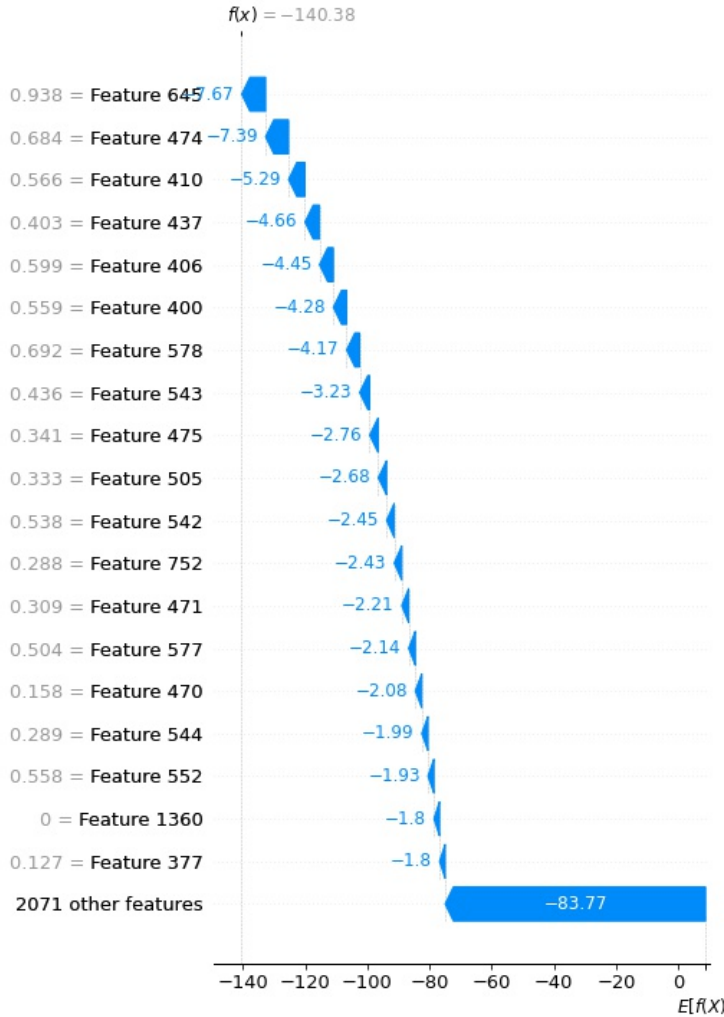
XAI, MLP:Z true pred 95.9 98.0



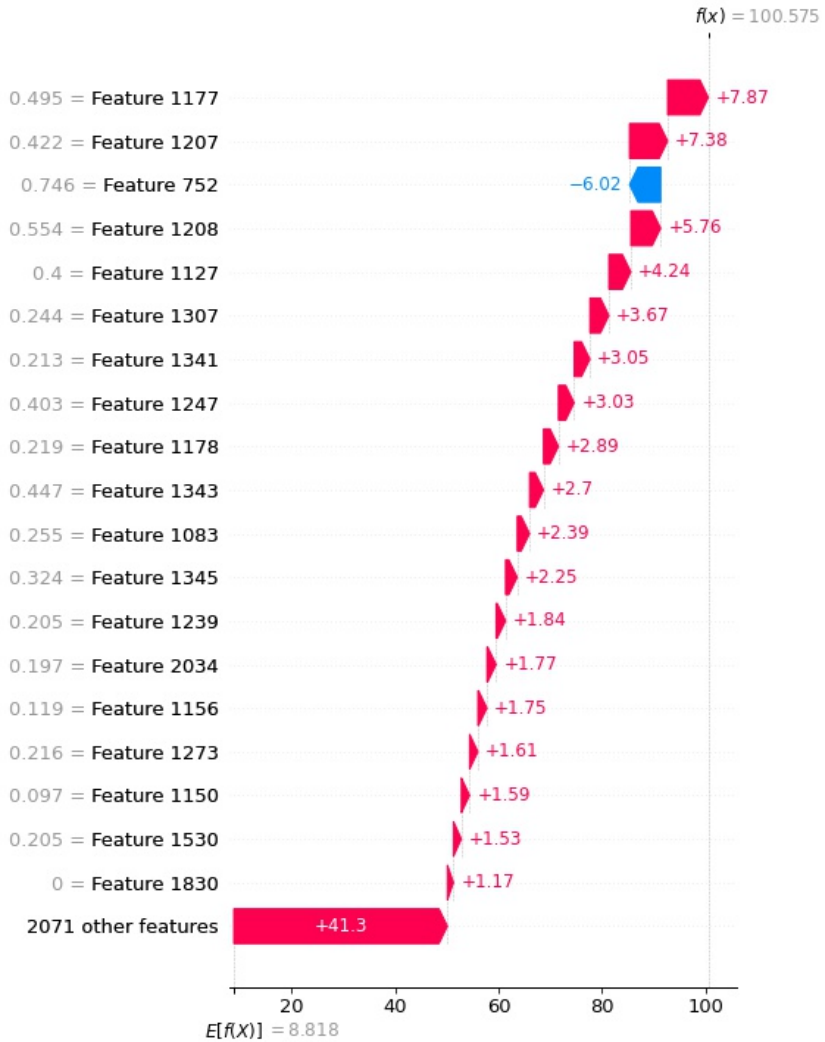


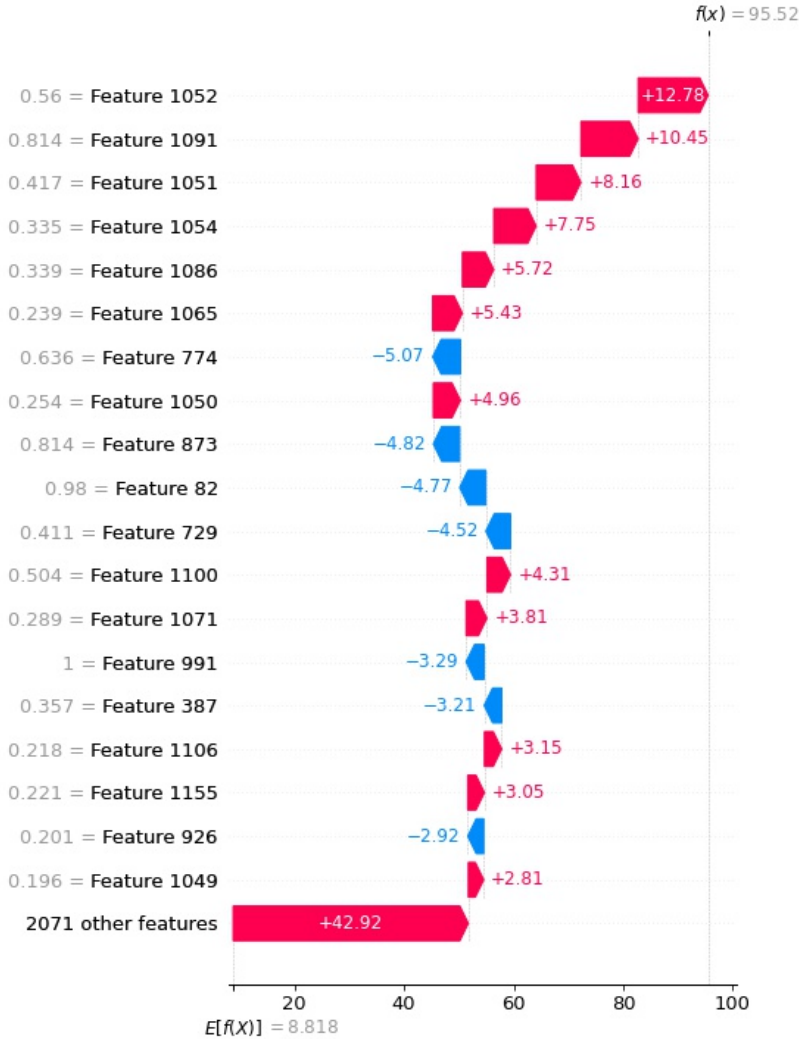
XAI, MLP:Z true pred -146.3 -140.4





XAI, MLP:Z true pred 140.7 100.6




f(x) = 95.52
0.56 = Feature 1052
0.814 = Feature 1091
0.417 = Feature 1051
0.335 = Feature 1054
0.339 = Feature 1086
0.239 = Feature 1065
0.636 = Feature 774
0.254 = Feature 1050
0.814 = Feature 873
0.98 = Feature 82
0.411 = Feature 729
0.504 = Feature 1100
0.289 = Feature 1071
1 = Feature 991
0.357 = Feature 387
0.218 = Feature 1106
0.221 = Feature 1155
0.201 = Feature 926
0.196 = Feature 1049
2071 other features
+12.78
+10.45
+8.16
+7.75
+5.72
+5.43
−5.07
+4.96
−4.82
−4.77
−4.52
+4.31
+3.81
−3.29
−3.21
+3.15
+3.05
−2.92
+2.81
+42.92
20
40
60
80
100
E[f(X)] = 8.818

[!ht]

Conclusion

Messages

- DL is a powerful tool, please use it smartly for **rare event research and underground physics**.
- DL is also in the frontier of knowledge, please don't use 20 years old algorithms (except if you need ultrafast algorithm for the trigger).
- XAI generates virtuous cycles, answering why NN/DL model produces a particular results, and arising new questions about how to improve (looking at error patterns).
- Don't scratch surface of ML, dig/dive in DL, and **ask to experts**.