Deep Learning applied to Capture Cross Section Data Analysis

A. Sánchez, D. Cano-Ott, E. Mendoza

CIEMAT– Unidad de Innovación Nuclear

Workshop on:

Machine Learning In Nuclear Science and Technology Applications

GOBIERNO DE ESPAÑA

MINISTERIO DE CIENCIA E INNOVACIÓN

Ciemat

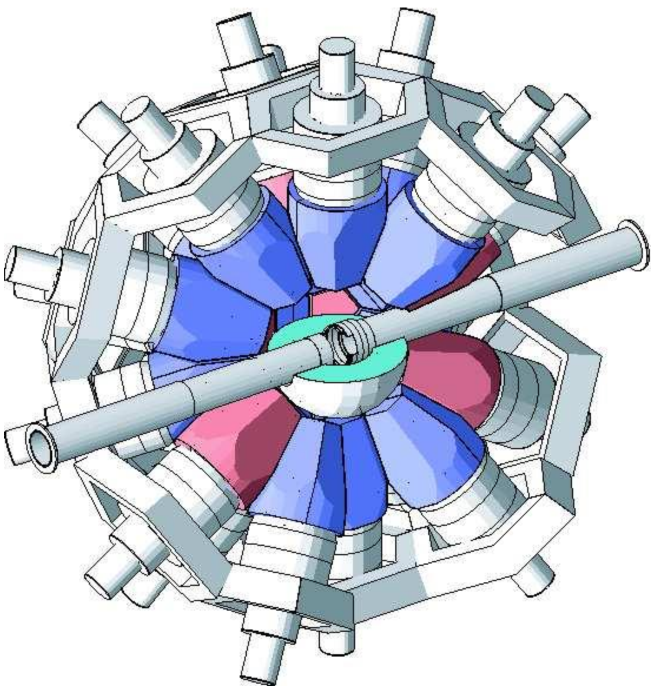
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Contents

DL and nuclear data: neutron capture classification

Experimental framework

- The **n_TOF Total Absorption Calorimeter** (**TAC**): a γ-ray total absorption detector for measurement of (n,γ) reactions cross sections.
- Main features:
  - Large solid angle
  - High efficiency
  - Good energy resolution
  - High segmentation
  - Low neutron sensitivity
  - Fast time response



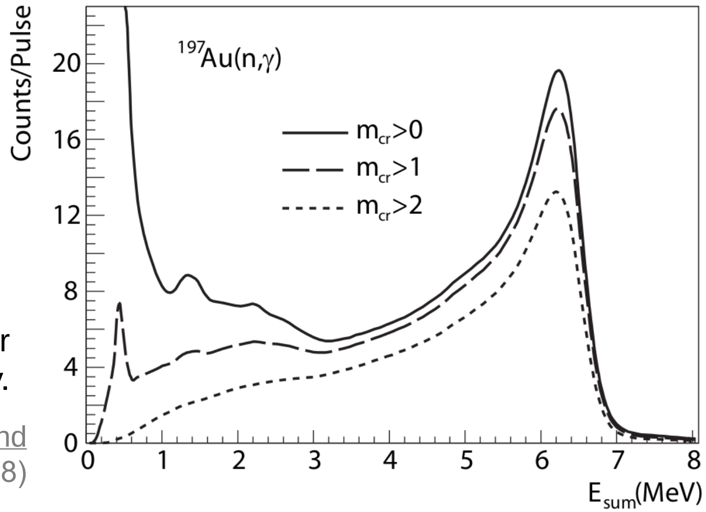
Half-opened 3D-model of the TAC
Image source: Carlos G. Guerrero Ph.D. thesis (2008)

Composed of **40 BaF_2 crystals**

DL and nuclear data: neutron capture classification

Problem definition

- The **Total Absorption Calorimeter** (**TAC**) at nTOF provides the deposited energy of gamma rays in 40 BaF_2 crystals.
- **Current method**: specific **cuts** or constraints on $\mathbf{E_{sum}}$ and $\mathbf{m_{cr}}$ values of each event to improve the signal-to-background ratio.
- **Objective**: to build a DL model for classification of capture/non-capture events from TAC signals and compare with current method.



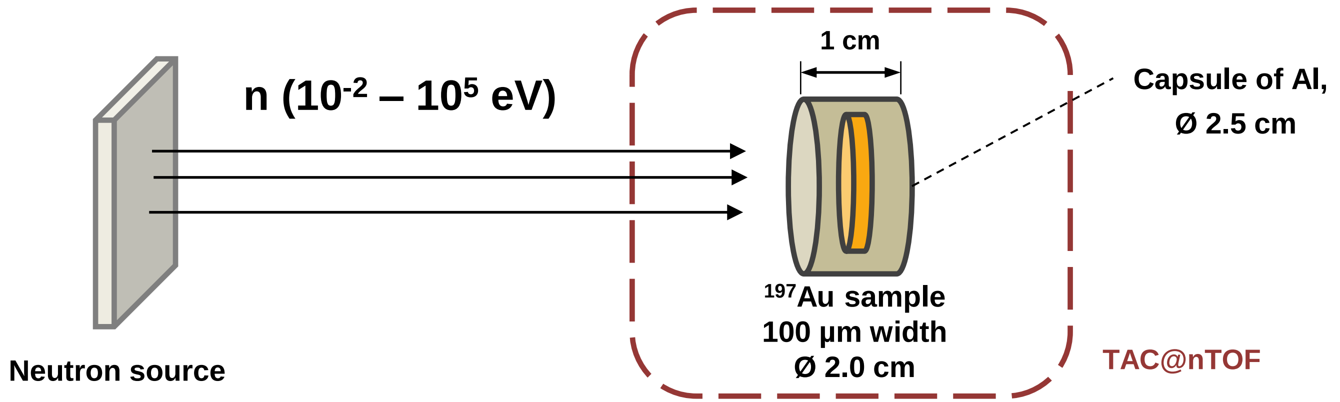
Deposited energy distribution of the 197Au(n,γ) measurement under different constraints in crystal multiplicity.

Image source: Carlos G. Guerrero Ph.D. thesis "Measurements of the 237Np and 240Pu neutron capture cross sections at the CERN nTOF facility" (2008)

DL and nuclear data: neutron capture classification

Methodology

- **Input**: 40-sized arrays of deposited energy in each crystal from **Monte Carlo simulations** of a **^{197}Au** measurement experiment.

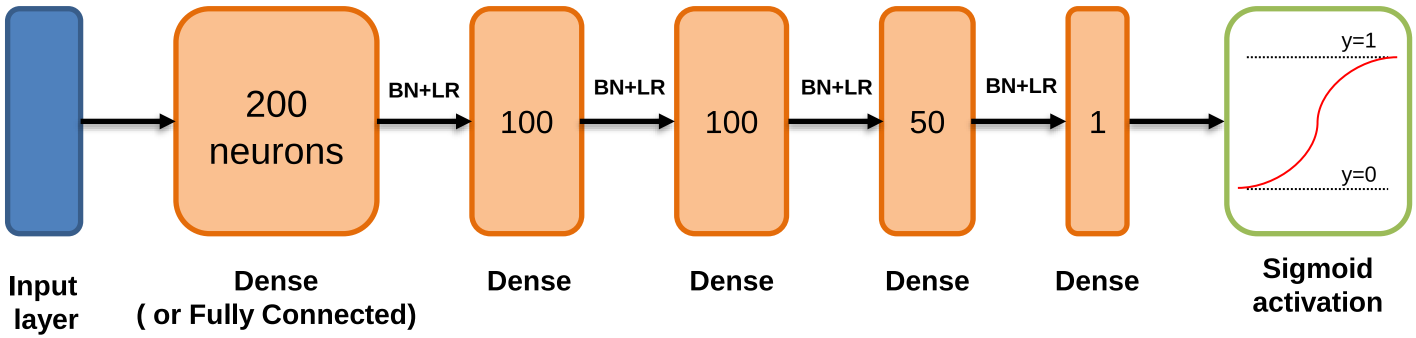


- **Neural Network classifier**: a model with 5 *dense* layers, 18K trainable parameters, binary cross-entropy loss function, and Adam optimizer, using the Keras API for Python. The NN classifies between capture (0) and non-capture (1).

DL and nuclear data: neutron capture classification

Methodology. Network architecture

BN+LR = Batch Normalization + Leaky ReLU activation



- Combinations of **DL with different cuts** on E_{sum} and m_{cr} are studied.
- The **performance** of all methods are **evaluated** through the values of:
  - $\varepsilon_{\gamma,rel}$ = capture detection efficiency. Set to 1 for the dataset without any cut.
  - R_i = signal-to-background ratio in the *i*-th interval of the incident neutron energy spectrum. Calculated as true capture events divided by non-capture events.

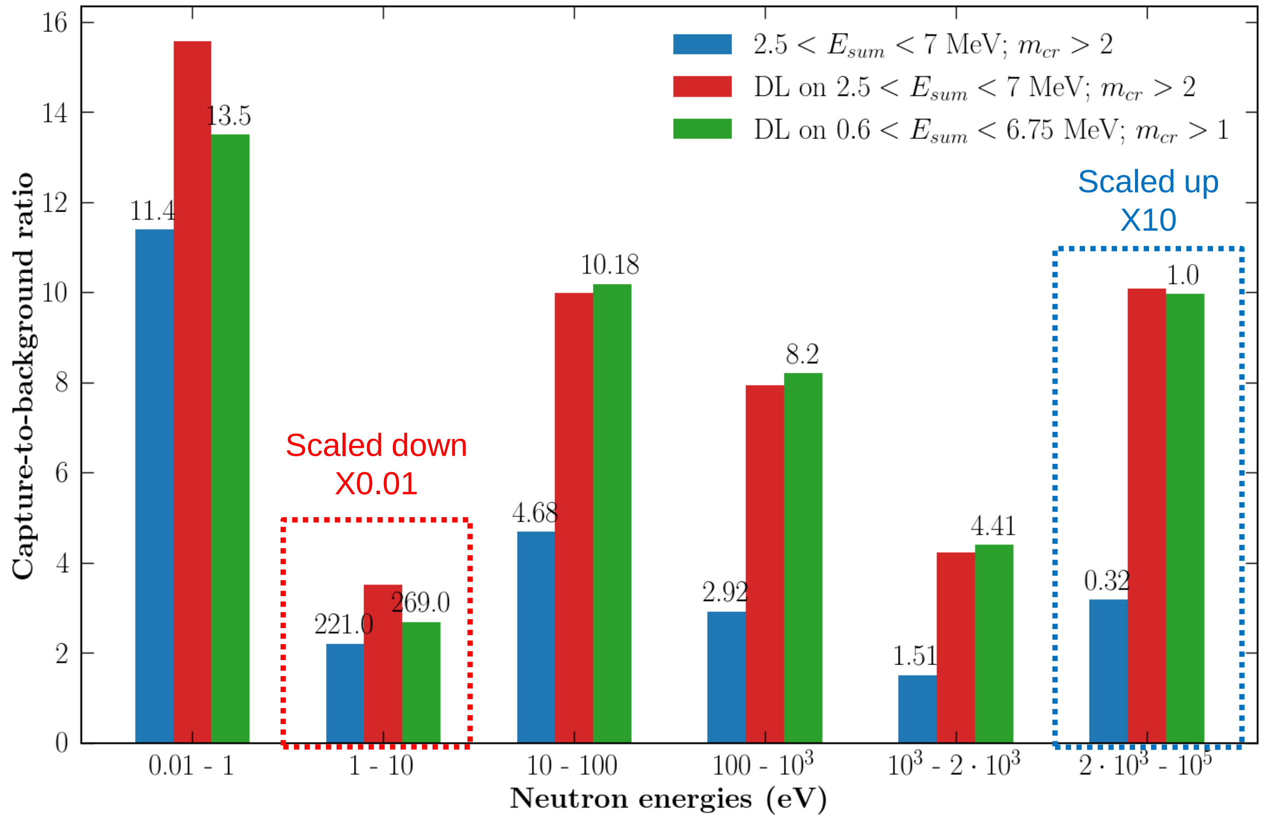
Results and comparison with traditional method

- We get **similar results** with **higher capture efficiency** by tuning the DL model parameters (# of layers, # of parameters, etc.) and by choosing more precise and relaxed cuts.

Capture efficiencies, $\varepsilon_{\gamma,rel}$:

- **Std. cuts** = 55%
- **Std. cuts + DL** = 43%
- **New cuts + DL** = **62%**

^{197}Au with aluminum canning



Results and comparison with traditional method

Case of a fissile sample: ^{239}Pu

- Example of using traditional cuts with or without DL for data including **fission events**.

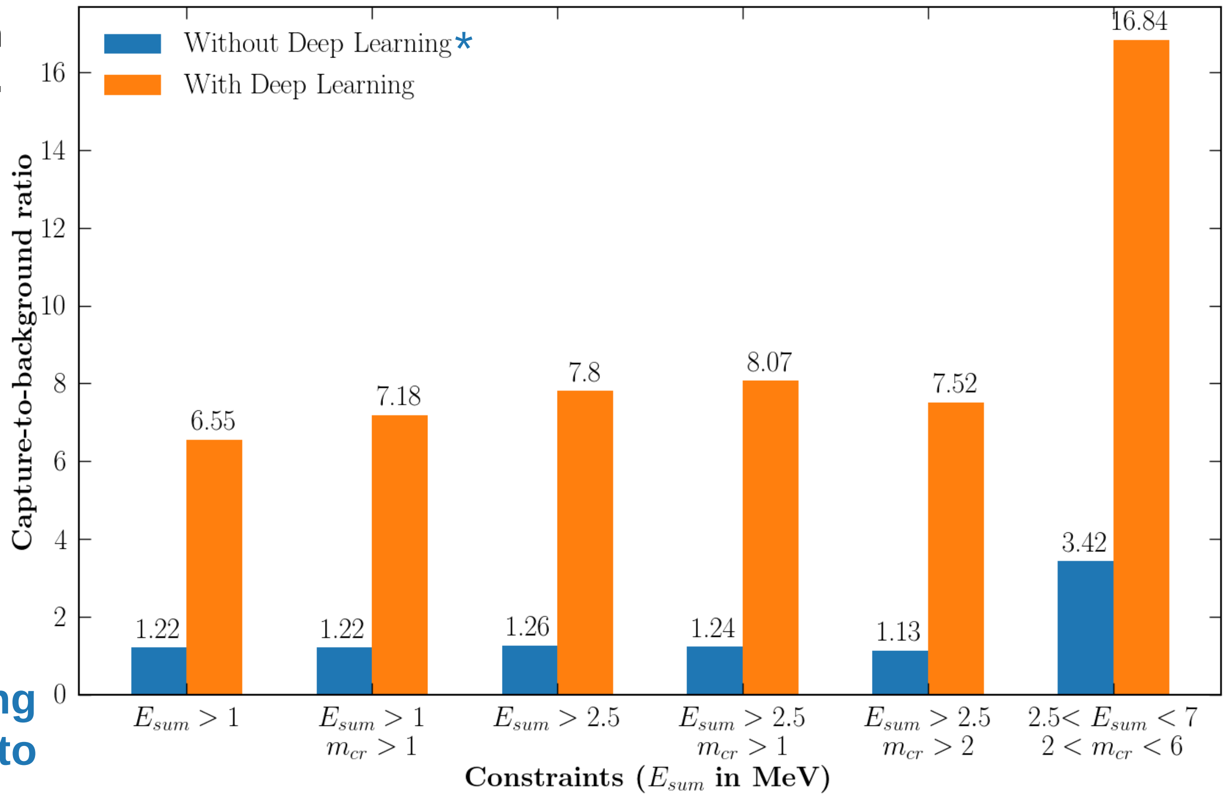
Figure shows the full-spectrum capture-to-background ratio.

Capture-to-fission rate in original data (no cuts): ~0.5

Some capture efficiencies, $\varepsilon_{\gamma,rel}$:

- **$2.5<E_{sum}<7;2<m_{cr}<6$** = 73%
- **Std. cuts + DL** = 63%
- **$E_{sum}>2.5;m_{cr}>1$ + DL**= 77%

(*): ratios calculated adding the $E_{sum}<7$ MeV constraint to the indicated cut (X axis).

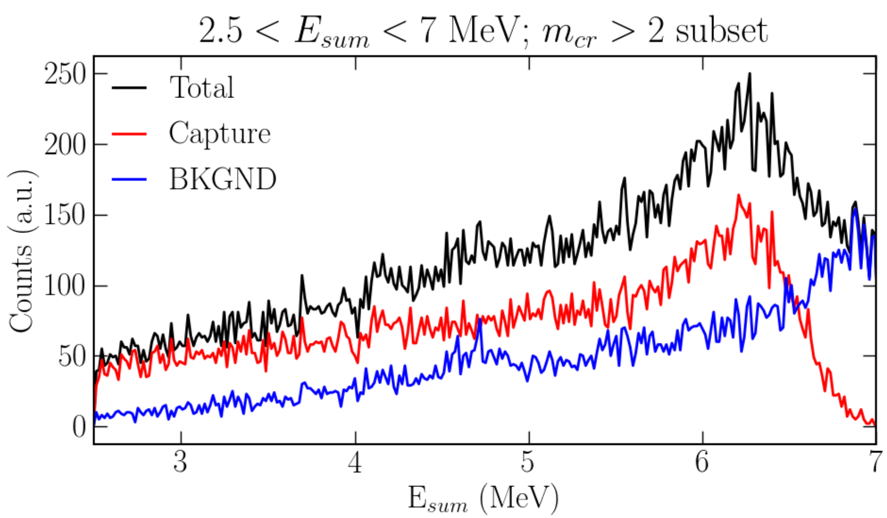


Results and comparison with traditional method

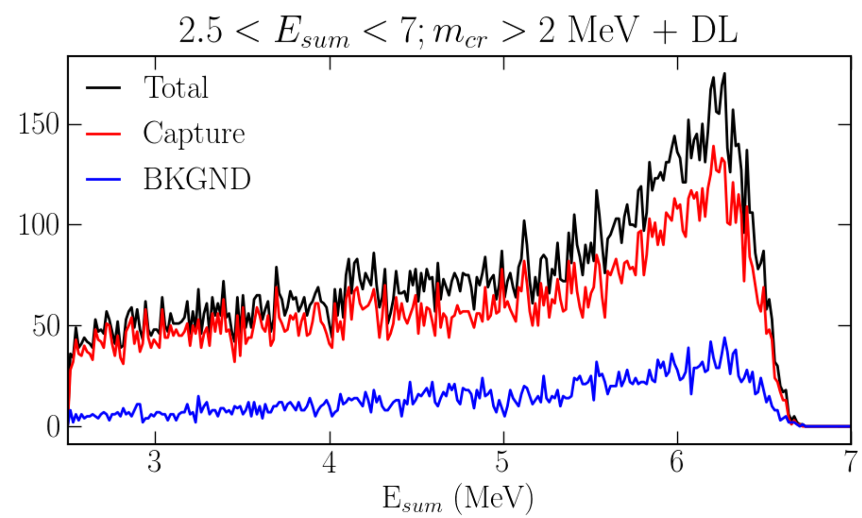
Deposited energy spectrum: ^{197}Au simulation

For neutron energies between 1 keV and 2 keV

Without Deep Learning



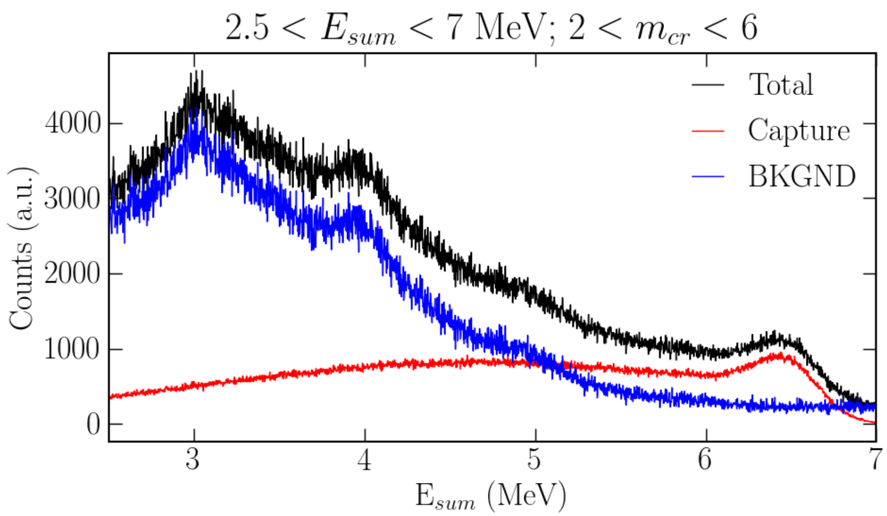
With Deep Learning



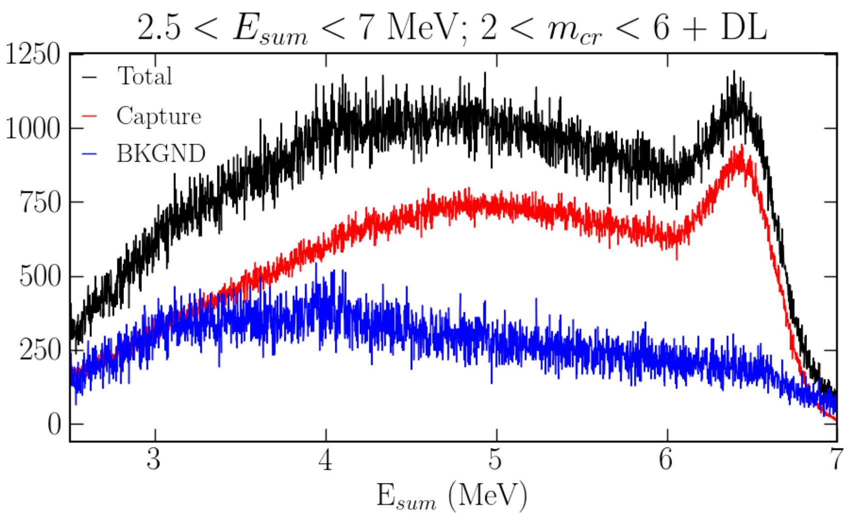
Results and comparison with traditional method

Deposited energy spectrum: ^{239}Pu simulation

Without Deep Learning



With Deep Learning

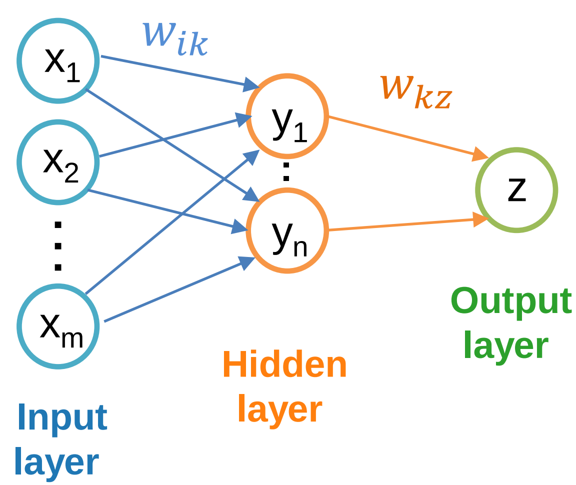


Additional results

Garson's Algorithm[1] for Variable Importance

- It estimates **relative importance** (**RI**) of **each variable of the input**, taking into account all the connecting weights between neurons from input to output.
- An attempt of **interpreting neural-networks** connection weights.
- For a NN with *m* input neurons (variables) and *n* hidden layer neurons, the RI of an input variable x_i is:

$$RI(x_i) = \sum_{k=1}^{n} \frac{|w_{ik} \cdot w_{kz}|}{\sum_{l=1}^{m} |w_{lk} \cdot w_{kz}|}$$



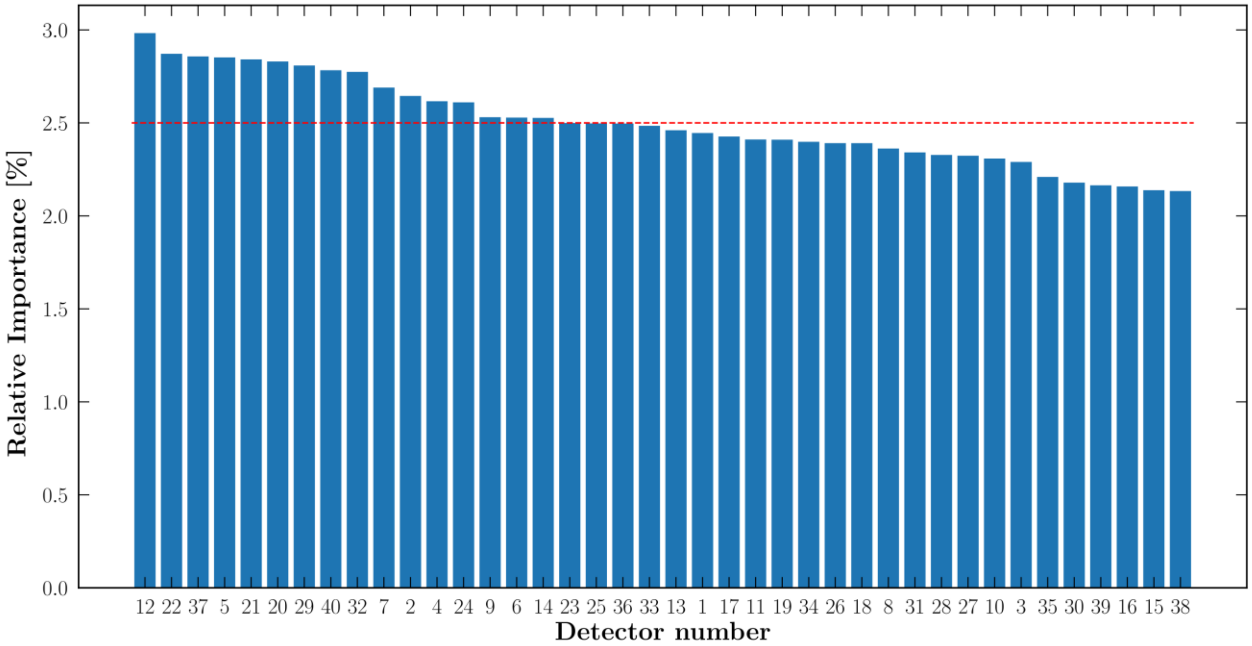
- **Limited** to Multi-Layer Perceptrons (MLPs) with **one hidden layer**.

[1]: Garson, G.D., *Interpreting neural-network connection weights* (1991).

Additional results

RI on the simulated ^{197}Au (n,γ) data with TAC

- We use a **very simple NN**: Dense(2000,ReLU)+Dense(1,sigmoid).



- Useful for checking the performance of individual crystals and, for example, for the design of new detectors.

Summary

- We have implemented **Deep Learning** techniques in the process of **capture event discrimination** of **simulated** data for a **^{197}Au sample with Al canning,** and a **^{239}Pu sample**.

- We find DL must be **used together with some cuts** to yield better results than the traditional method.

- In the particular case of the **^{197}Au sample**, by using DL we can obtain **signal-to-background ratios ~3 times larger** than traditional cuts $\{2.5<E_{sum}<7,\ m_{cr}>2\}$, with a **~13% larger capture efficiency**.

- For the case of the **fissile sample ^{239}Pu**, using the std. cuts + DL yield data with a **ratio ~5 times larger** than using only the std. cuts, with a ~14% reduction in capture efficiency. By using *softer* cuts + DL, we get ~2 times larger ratios and ~16% larger efficiencies.

- We also explored the use of Garson's Algorithm of Variable Importance.

Conclusions

- **Deep Learning** can be an interesting tool for nuclear data analysis that seems to **improve the performance of prior methods**.
- However, Deep Learning is still, to a great extent, a **trial and error method**, and therefore **time-consuming**.
- In addition, Deep Learning is probably problem-dependent. Difficult to generalize.

Future work

- Test the methods with real experimental data and apply to other cases.
- Further study of methods similar to Garson's algorithm in order to better understand the performance of neural networks.

Questions?

THANK YOU!

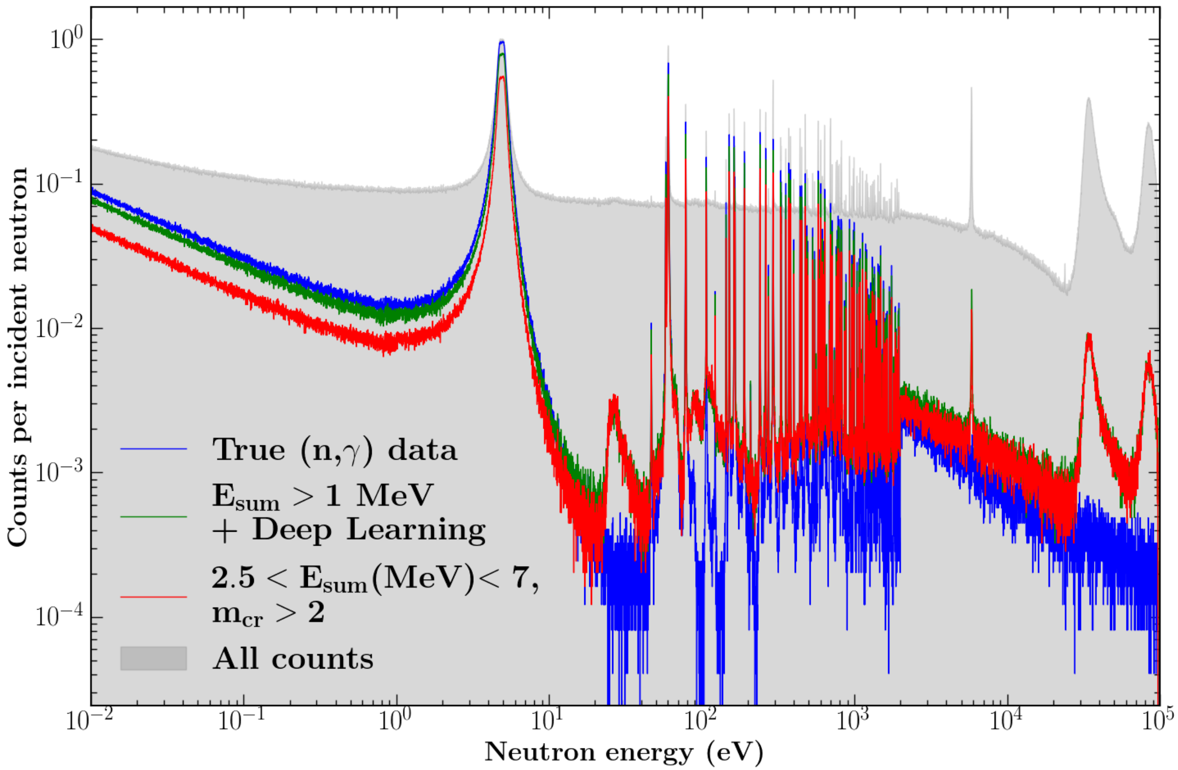
Results and comparison with traditional method

Example of capture spectra with and without DL

A **combination of DL with lax cuts** discriminates most of the capture events, **similar to the results of the common strong cuts**.

In addition, it improves capture efficiency in some regions.

^{197}Au with aluminum canning



Results and comparison with traditional method

Effect of using some cuts before DL

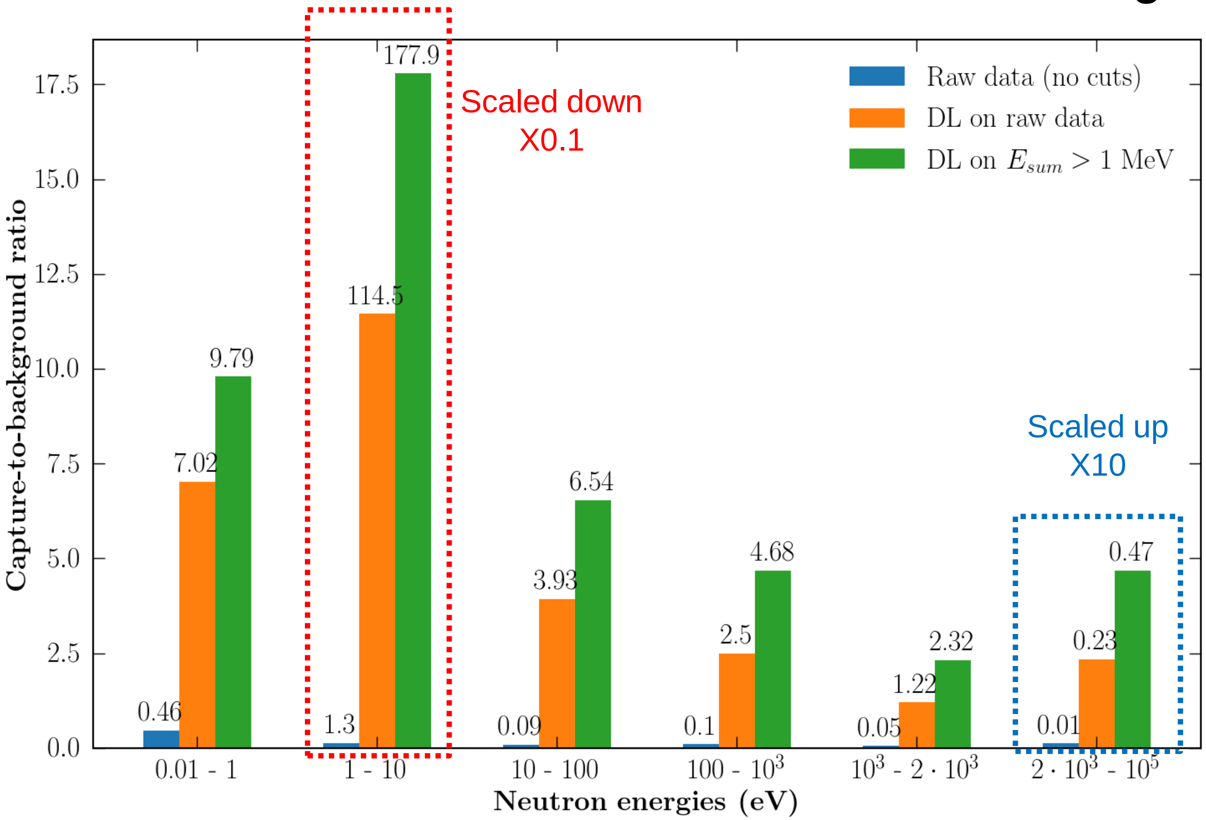
To use **only DL** clearly improves raw data ratios, but **not enough** to compete against traditional cuts.

Combining cuts + DL gets better results.

Capture efficiencies, $\varepsilon_{\gamma,rel}$:

- **Raw data** = 100% (by definition)
- **Only DL** = 93%
- **E_{sum}>1 + DL** = 81%

^{197}Au with aluminum canning



Results and comparison with traditional method

DL with more restrictive cuts

Capture efficiencies, $\varepsilon_{\gamma,rel}$:

- **Std. cuts** = 55%
- **Cuts#1+DL** = 81%
- **Cuts#2 + DL** = 77%
- **Std. cuts + DL** = 43%

Above 10 eV, the cuts+DL methods improve the ratio of the std. cuts.

Best results are achieved using **std. cuts + DL**, obtaining an improvement of a **~2-3 factor,** with the cost of reduced efficiency.

^{197}Au with aluminum canning