ATRAPP Meeting

Thesis Planning

Jeremy Couthures

Official ITk plots

1) **Move the Hashing code to core:**
    1) Move the code to core (local) √
    2) Make the code compile √
    3) Ensure same results than before √
    4) Pull request
2) **Link ACTS+Hashing version in Athena (TWiki) √**
3) **Edit seeding tool (in Athena) with Hashing**
4) **Reproduce official plots (with Florencia):**

    Athena:
    1) Default
    2) ACTS
    3) ACTS SSS only
    4) ACTS PPP only
5) **Reproduce the plots with Hashing (Eff + CPU)**

Next

1) Get the data from GNN people (SP and truth) (Alexis) (Dumped files) (Dumping code)

2) Train NN (using acorn? It can already read the dumped files)

Planning (Part 1)

1. Data:

- 1: Redo the official ITk plots inside Athena (03/04/2024)
- Plan A: Use only the space points and the associated truth particles from GNN people (05/04/2024)
- Plan B: Use Acorn (A Charged Object Reconstruction Network) (gitlab) with Athena samples (12/04/2024)

2. Architecture:

- Plan A: Define the architecture for the model and be able to train it (12/04/2024)

3. Training: (Try to get as much details as possible: learning curve, hyper parameters, ...)

- Plan A: Have a model with "good performances" (to be defined) (26/04/2024)

Planning (Part 2)

4. Model analysis:

- Plan A: Identification of relevant neurons through the analysis of their activation:
  - variation of the values of the activations on the data → no variation = not a variable (= constant) (correspond to the dependency of the neuron with respect to the input values) (Neural activation pattern) (10/05/2024)
  - Permutation Importance on the activations of the neurons to evaluate the impact on the prediction (score) (31/05/2024)
- Plan AB: Direct analysis of the neurons:
  - Network Dissection method (correlation between the values of the neurons and physics variables)
  - Symbolic regression with pySR on the output neurons (and relevant neurons if found) and the relation with physics variables
- Plan ABC: Global analysis of the model:
  - Identification of the data-flow inside the model
  - Search for symmetries inside the network

5. Result analysis:

- What worked, what did not? (list)
- Issues, potential solutions
- Relevance of the results (Are they useful?) (do they bring something?) (can they be used?)