



~~ML~~ Lunch A.I.

Problem of the (almost) quarter

Thomas Britton
David Lawrence

Sept 2019

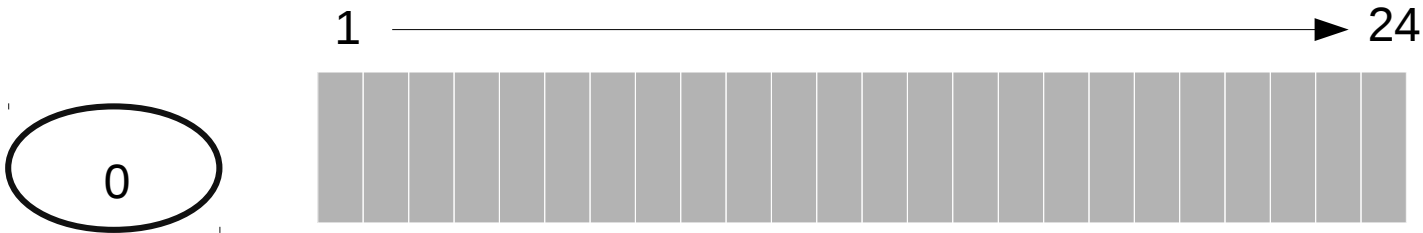
The GlueX FDC

- A tracking detector comprised of 4 “packages” of 6 wire planes
 - A total of 24 measurement planes
- Of interest to tracking is the state vector of a particle at the planes



The state vector

- 6 parameters
 - 3 for position
 - x, y, z
 - 3 for momentum
 - p_x, p_y, p_z
- We want this vector at a given plane
 - Always start from target (plane 0)
 - Magnetic field effects position
 - Energy loss (material) effects momentum
- Total number of parameters: 150



Notable difference

- Participants are given a data set that includes the complete state vector at every location. It will be up to participants to utilize the data appropriately
 - **You** break up the data into sets
 - **You** write your training script
 - **You** write your test script
- Do keep in mind the test data will have a fixed format (see next slide)

The goal

- We desire a projection to a given plane (z-value)
 - 5 parameters
 - x, y, p_x, p_y, p_z
- So given n state vectors (target always given) we would like the state vector of $n+1$ (z is given so the machine need only predict the remaining 5 parameters)

Test format

- $6n+1$ input parameters

- For the project to layer 2 ($n=2$) (target+1 layer given)

- $X^0, y^0, z^0, p_x^0, p_y^0, p_z^0 \left| x^1, y^1, z^1, p_x^1, p_y^1, p_z^1 \right| z^2$

- The expected output for the above track would be a set of 5 parameters (we will always expect 5)

- $x^2, y^2, p_x^2, p_y^2, p_z^2$

What is due?

- Given the flexibility on what you do with the data we expect all scripts you used for training **and** testing.
 - Wait and testing?!
- We also expect a single csv file with one line per track. Each line will contain 5 parameters in the order x, y, p_x, p_y, p_z

FAQ

- So I can use pyTorch?!
 - Yes
- So I can write my own tensorflow to do the problem in [LoLCode](#)?
 - Why would you do this to yourself?....but also yes
- So I can pre-process the data as I see fit?
 - Yep!
- I can cheat and use a standard algorithm
 - no

Bottom line

- You must use machine learning
 - We should be able to run your scripts and reproduce your results (you need to turn them in)
 - We reserve the right to reproduce any result.
- Non-ML solutions will be disqualified
- On the due date a test will be released. The “labels” will be withheld. The submission will be due **24 hours** from the time the test set is released

When will the test set be released?

- **October 30th by noon**
- The link to the needed files can be found here:
<https://halldweb.jlab.org/talks/MLlunch/Sep2019/>