5D Calorimetry Meeting

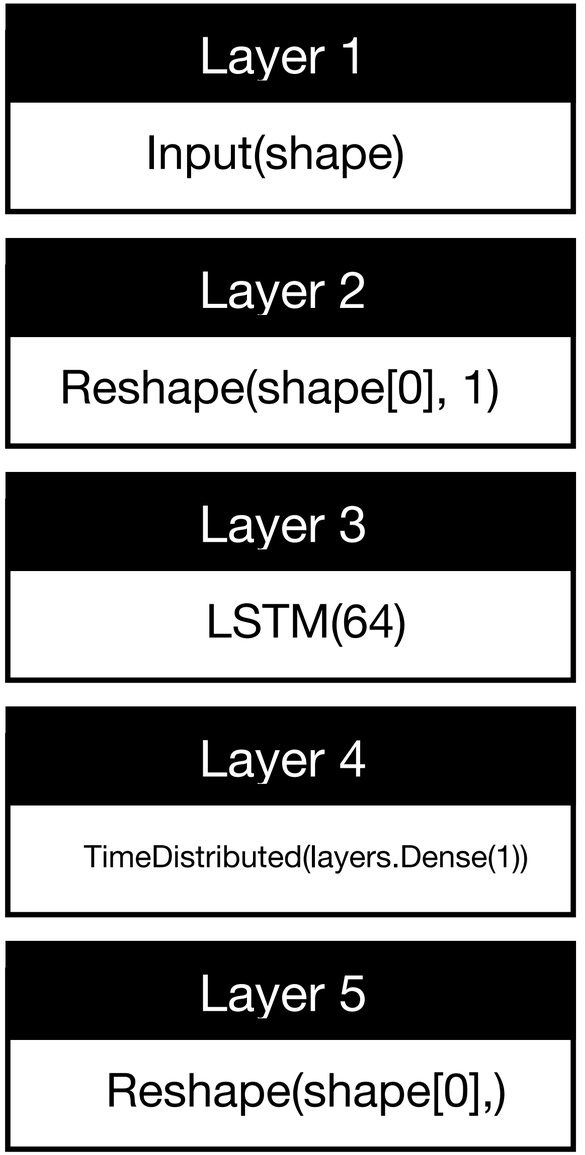
ML for pulse recovery

H. Brittan - 06/25/25

LSTM Model

Functionality

- Synthetic Pulse Generation with noisy landau pulse convoluted over truth “photon timing”, with 50,000 events
- This is a sequence-to-sequence LSTM model designed to process sequential data and output a value for each time step
- It is suitable for tasks like time-series forecasting, signal reconstruction, or sequence modeling where temporal dependencies are important



Data Report

Convoluted Photons in Muon Sim

Input File: mu20000evt_sim_pulse_data.h5

Samples: 100000

Time bins per sample: 750

Time per bin: 0.0400 ns

Total time window: 30.00 ns

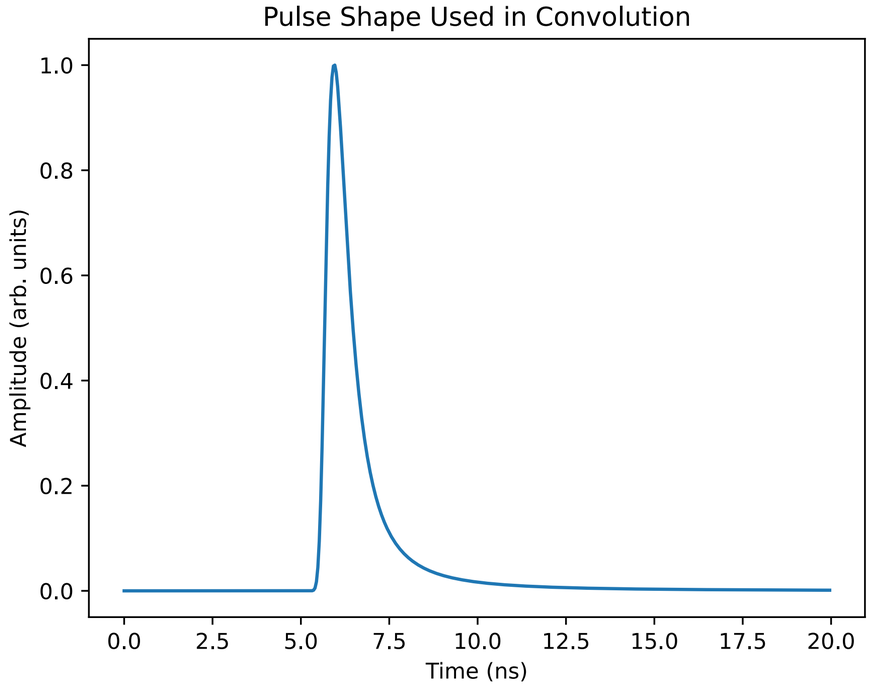
Avg. photon hits/sample: 14.51

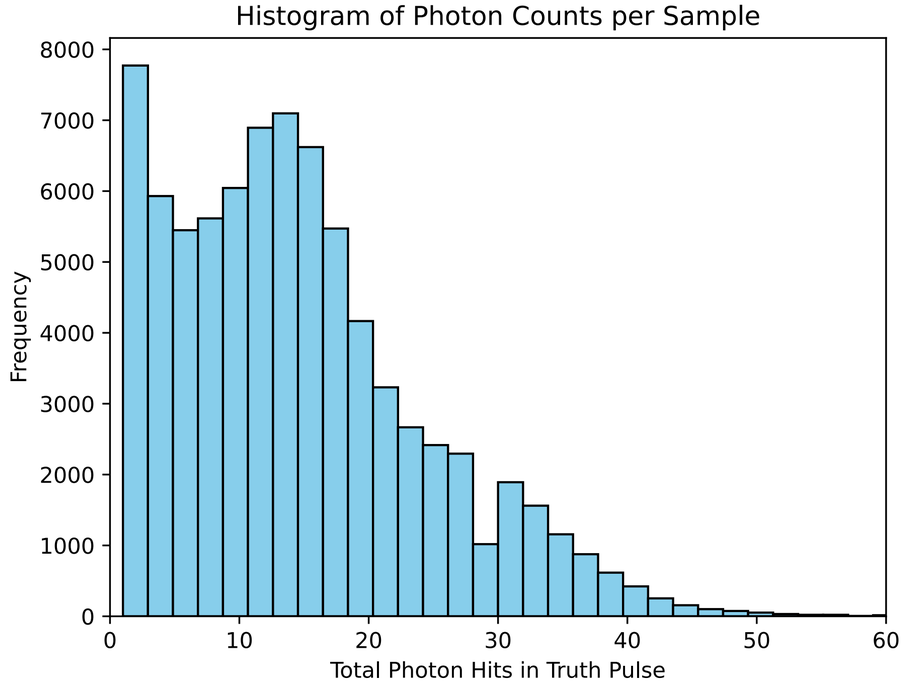
Avg. signal peak height: 12.34

Max. signal peak height: 95.56

Samples with hits: 100000

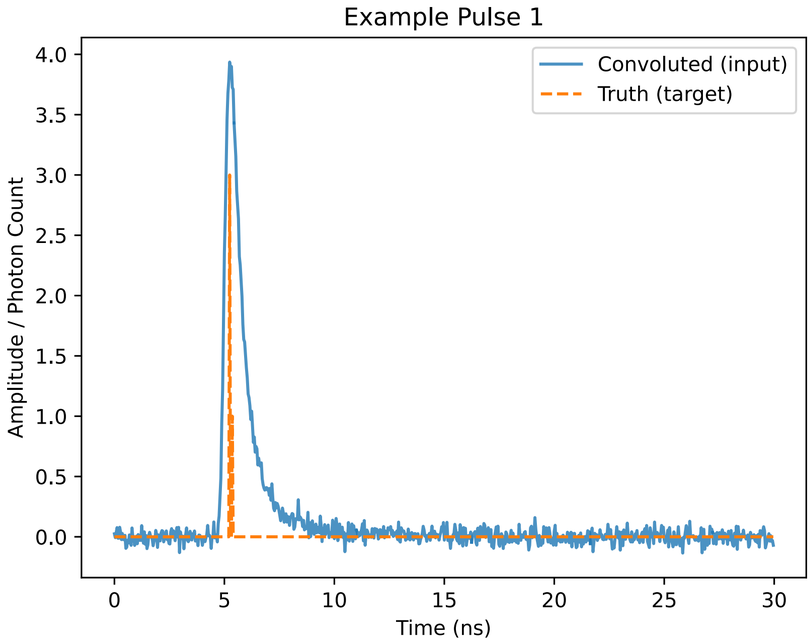
Pulse FWHM: 0.76 ns

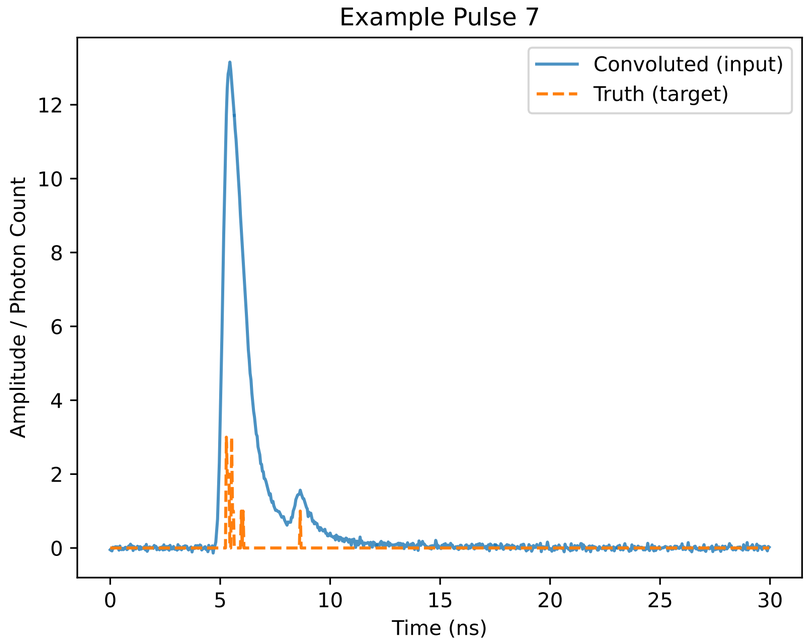


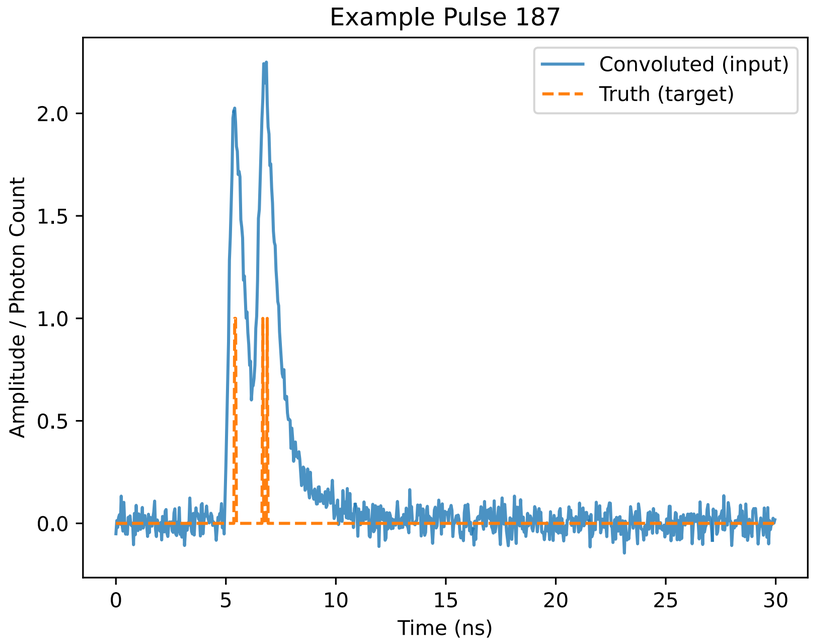


Pulse Visualization

Input data



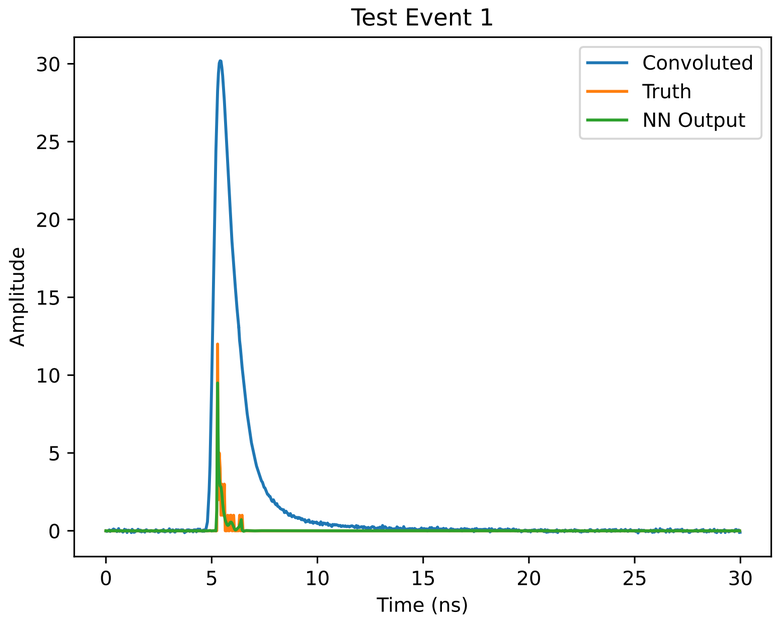


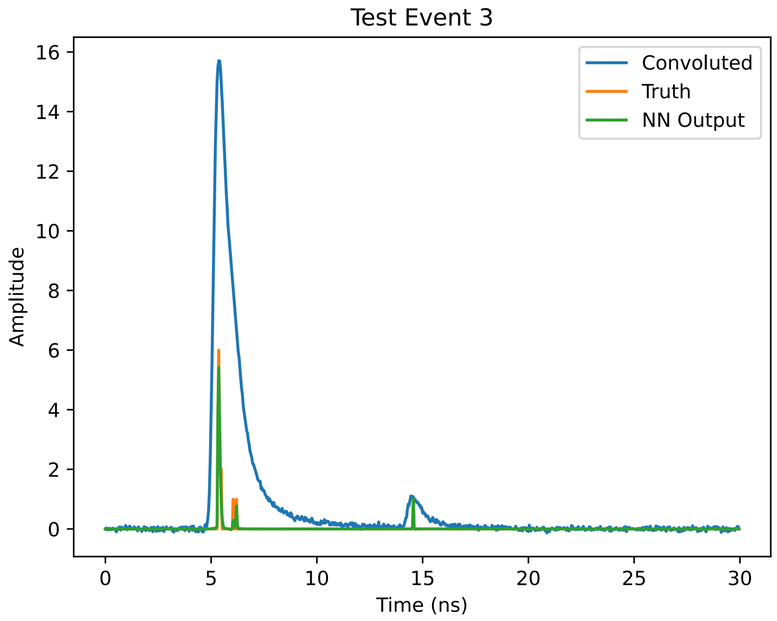


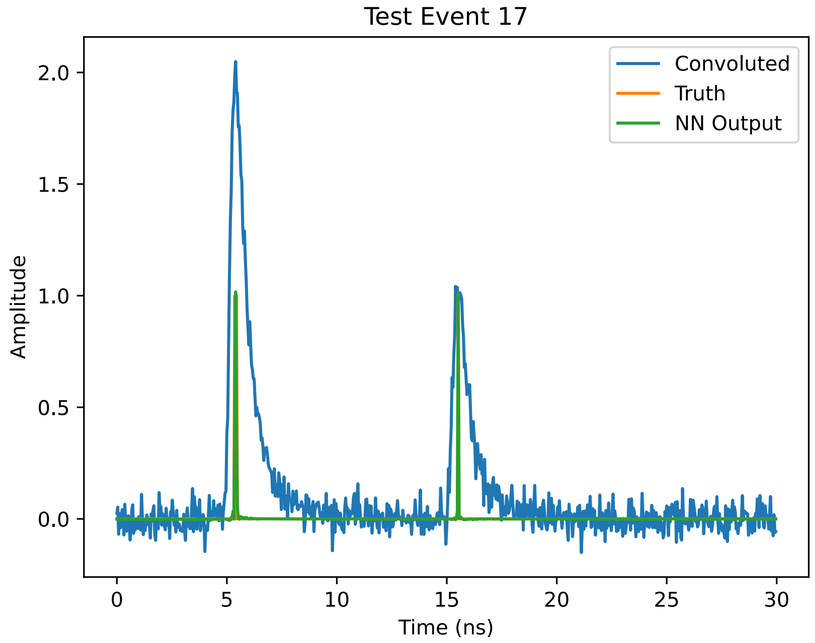
Examples of truth photon timing and their corresponding convoluted pulses

Pulse Visualization

output data (no shift)



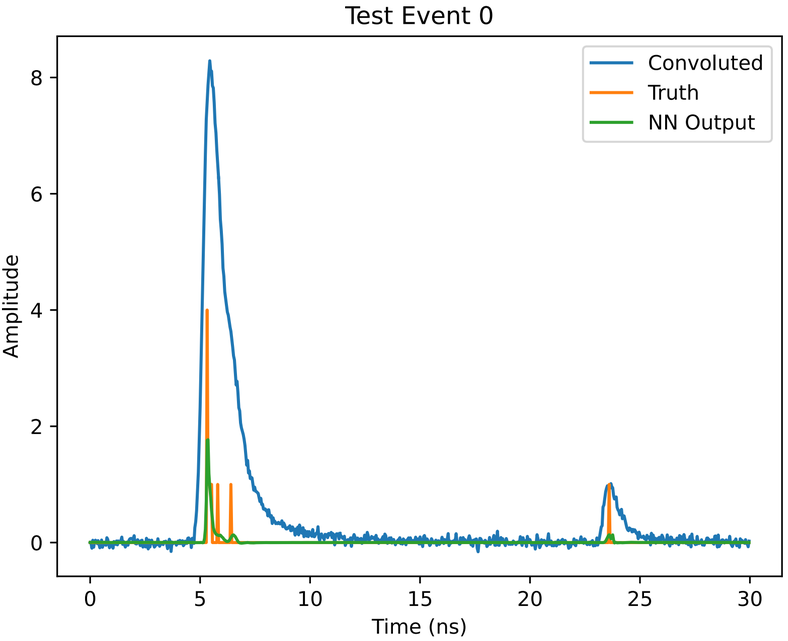


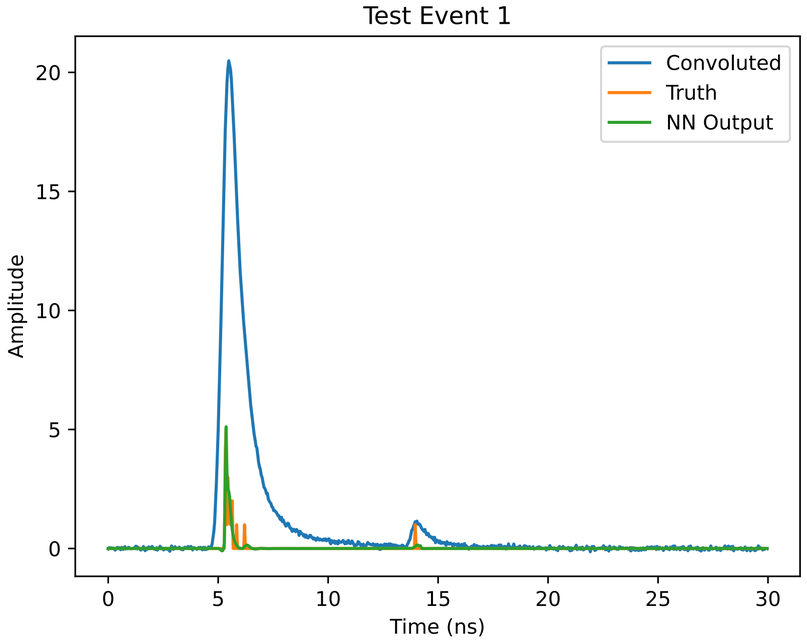


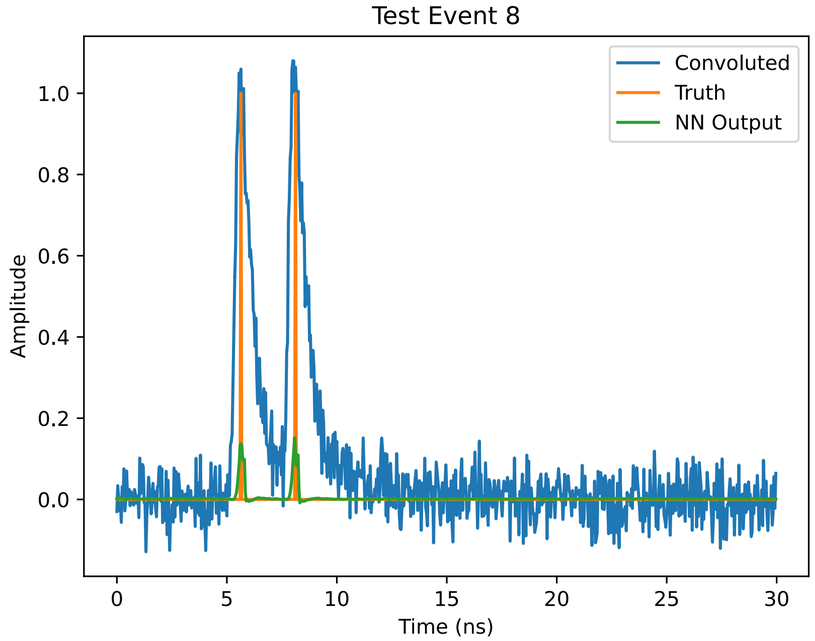
- Confidence with no shift evident in the smaller pulses after large main pulse
- Characteristic of LSTM model for learning index of truth value

Pulse Visualization

output data (± 0.2 shift)



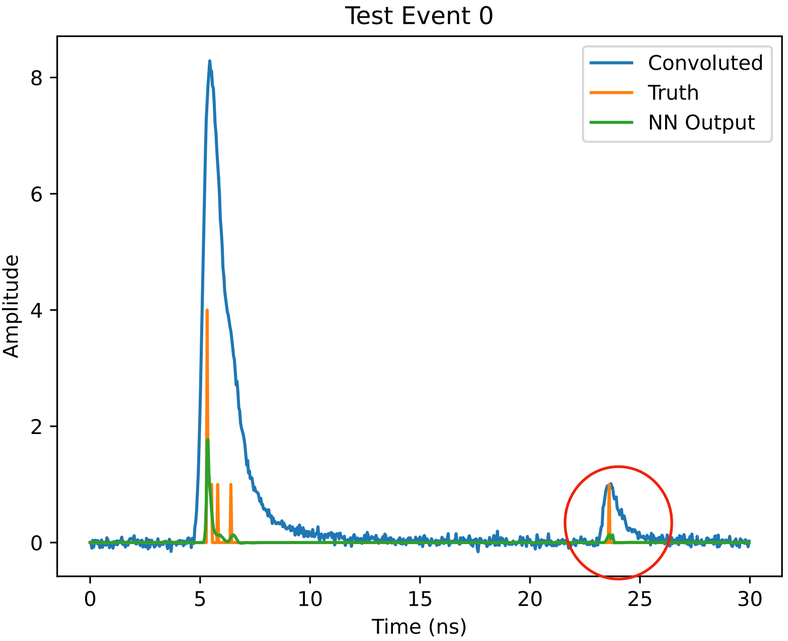


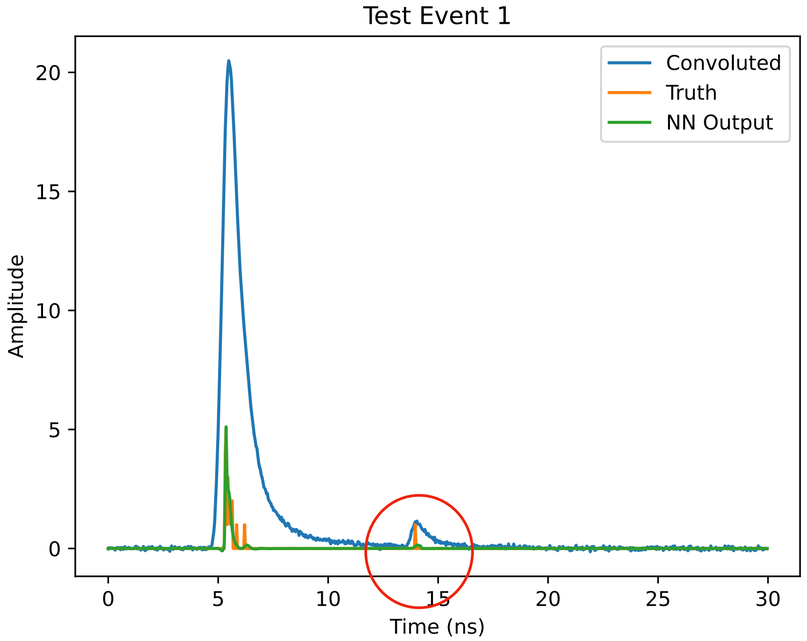


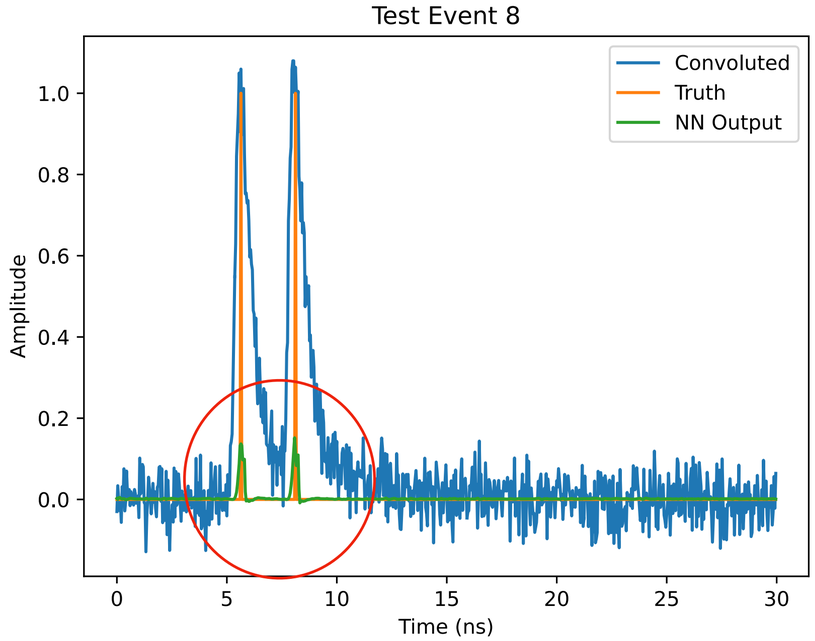
- Confidence with no shift evident in the smaller pulses after large main pulse
- Missing pulses and amplitude because of the changing index

Pulse Visualization

output data (± 0.2 shift)





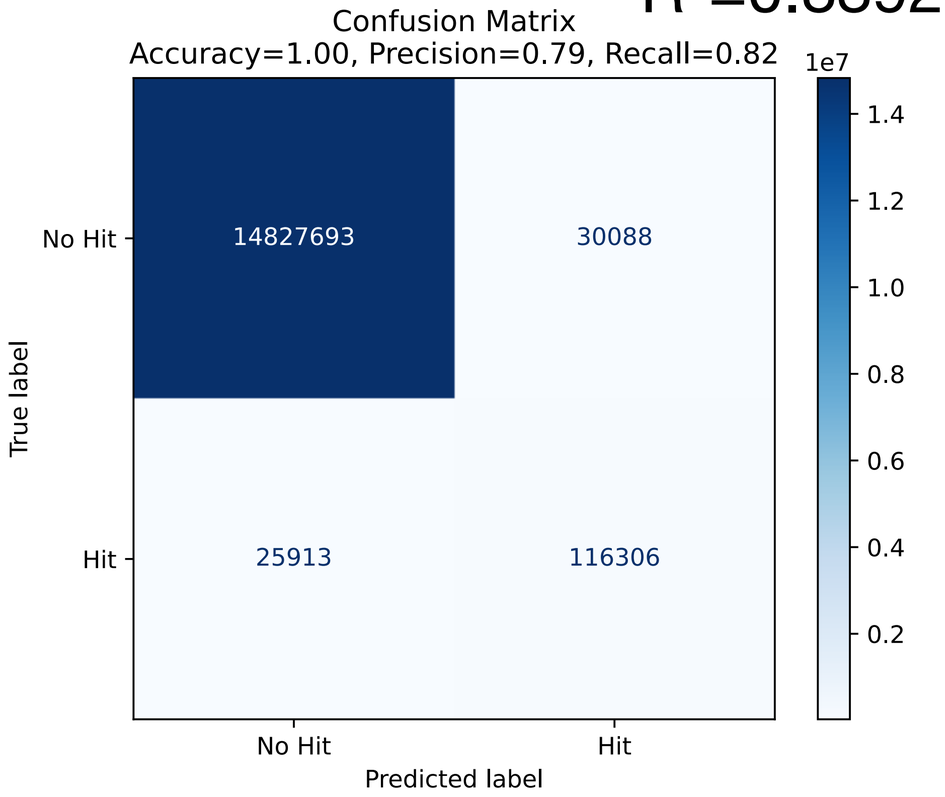


- Confidence with no shift evident in the smaller pulses after large main pulse
- Missing pulses and amplitude because of the changing index

Performance

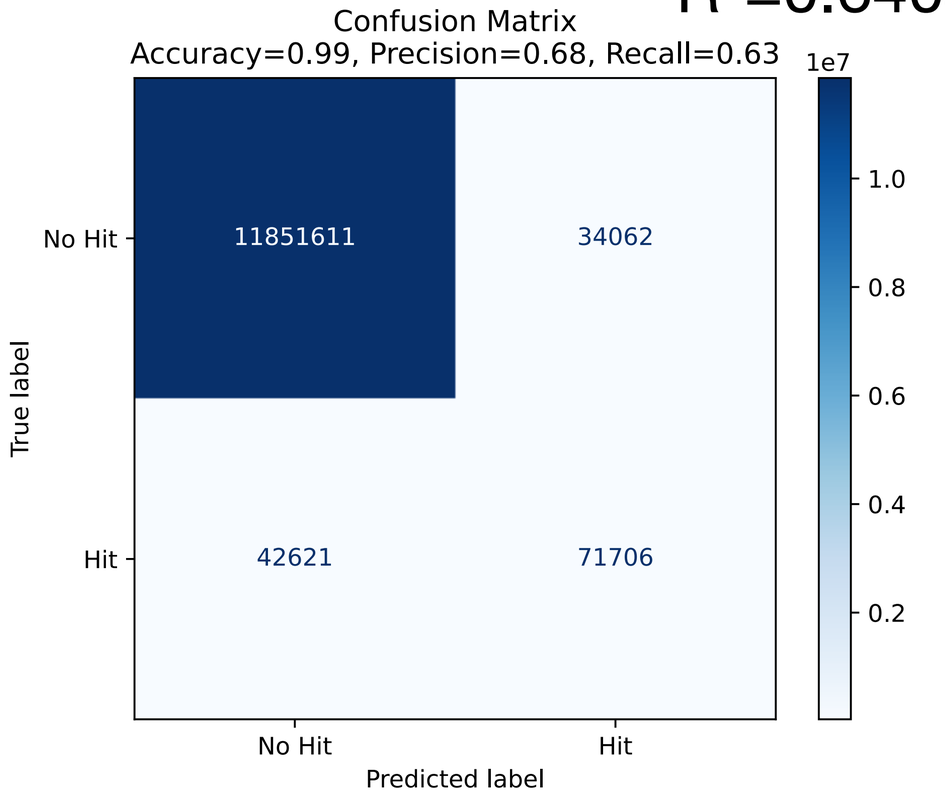
LSTM (no shift)

R^2=0.8392



LSTM (±0.2 shift)

R^2=0.6403



$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Next Steps

ML for pulse recovery

- Evaluate 1d CNN performance on random shift
- Consider other models to better account for shift and not learning truth index
  - CNN-LSTM hybrid
- Move to testing on cosmics data