Machine Learning-based Waveform Analysis for Pileup Events

Laxman Paudel

CoSSURF

15th May 2024

UNIVERSITY OF SOUTH DAKOTA

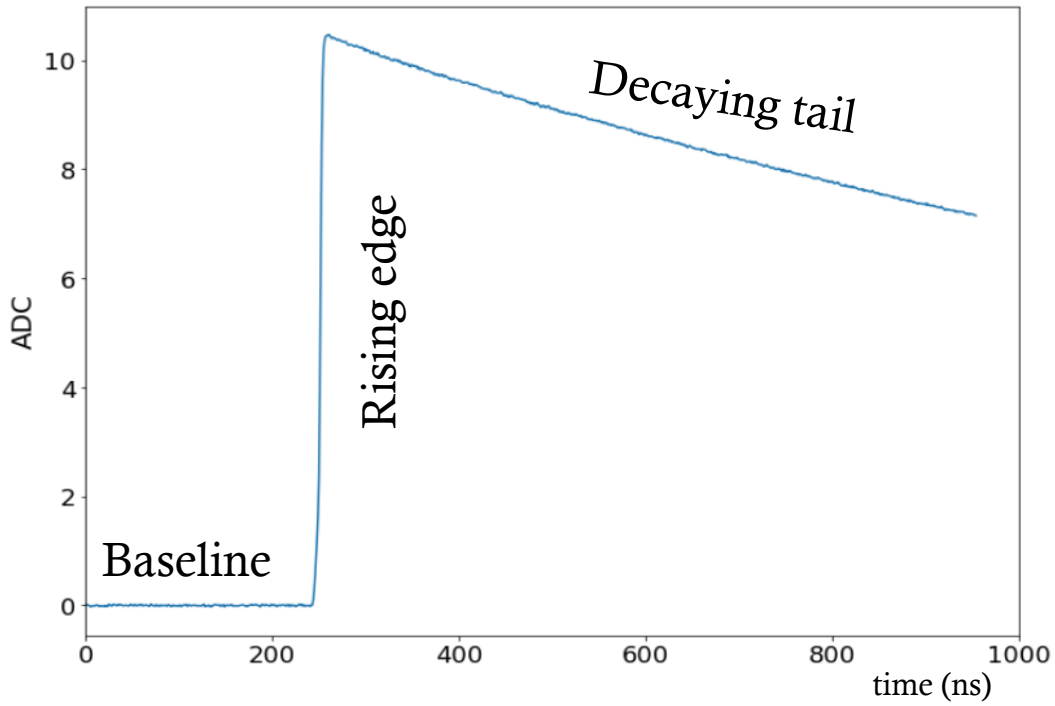
Outline

- Pileup waveforms
- Dataset creation
- RNN network implementation
- Network classification performance
- Network regression performance
- Summary and outlook

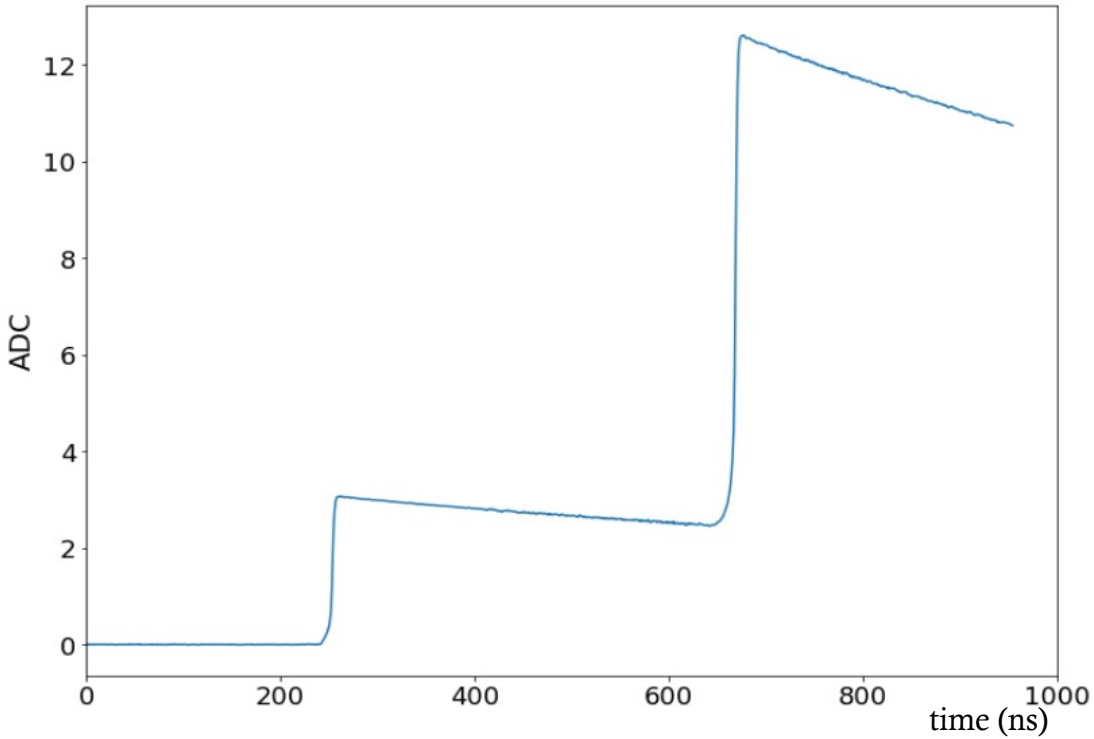
Pileup waveforms

- Pulses can pileup in an event
- Random pileup
- Gamma cascade can also create pileup events
- Isomeric transition, multisite events

Single waveform

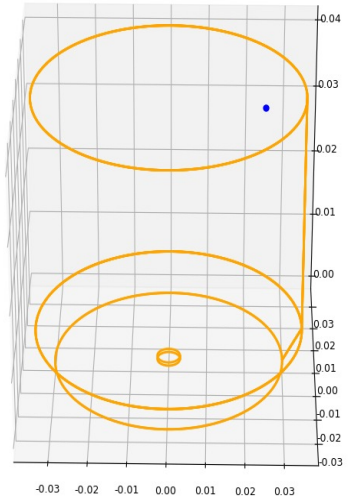


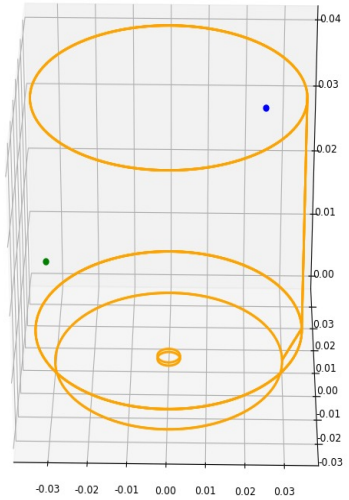
Pileup waveform

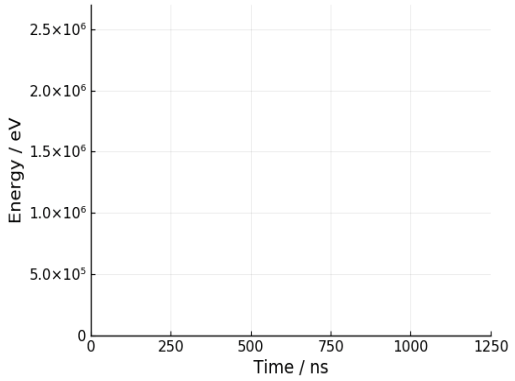


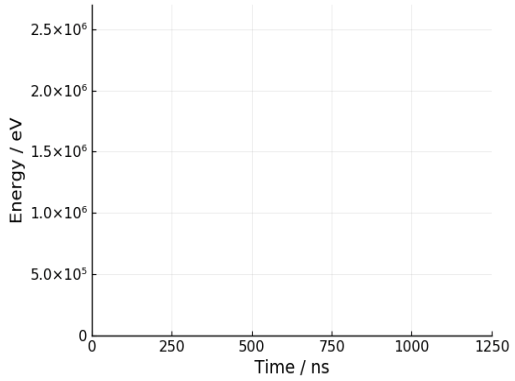
Motivation for machine learning-based waveform study

- Distinguish between pileup and single waveforms
- Machine learning based pulse shape analysis has a potential to outperform traditional analysis
- We can simultaneously train all detectors to avoid detector by detector tuning
  - This will be important for next-generation experiment that uses a larger number of channels like in LEGEND experiment.





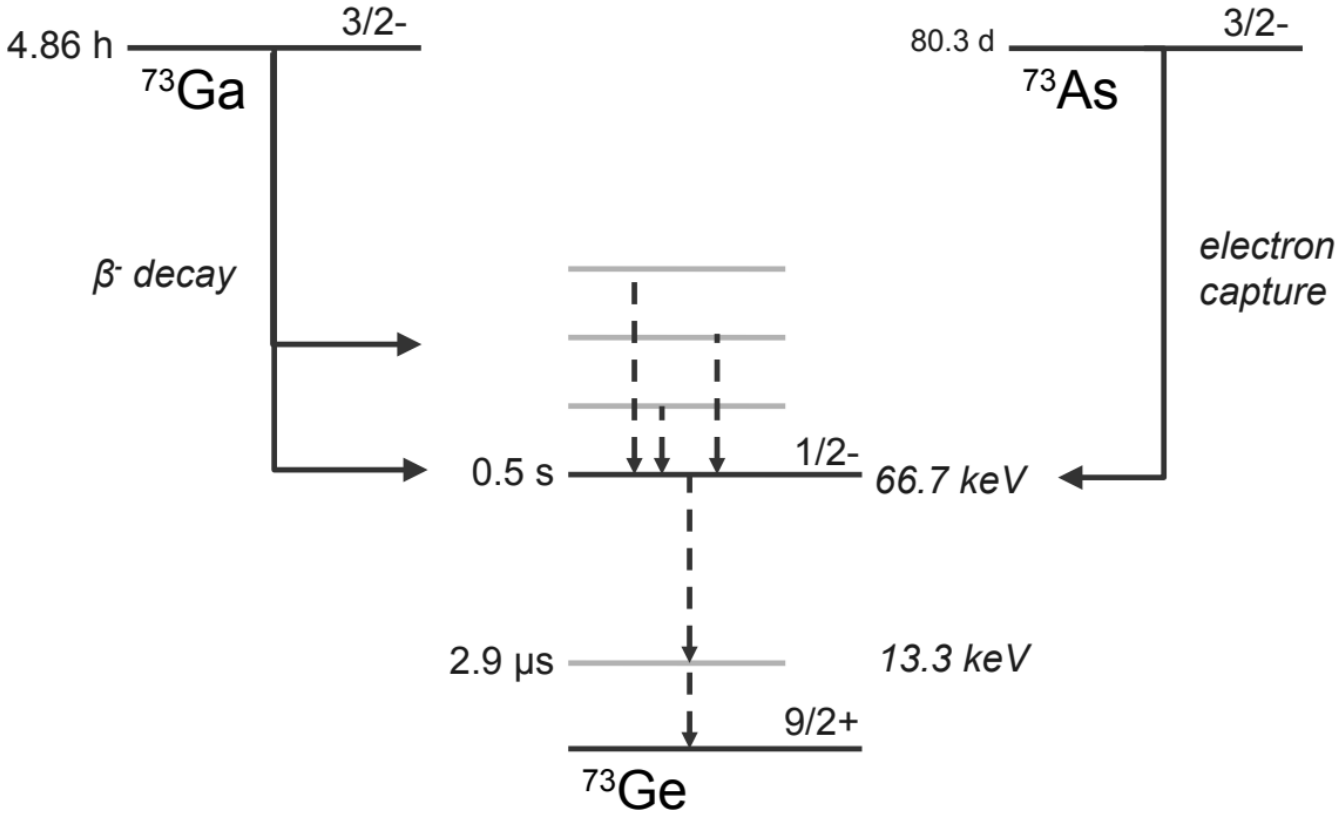


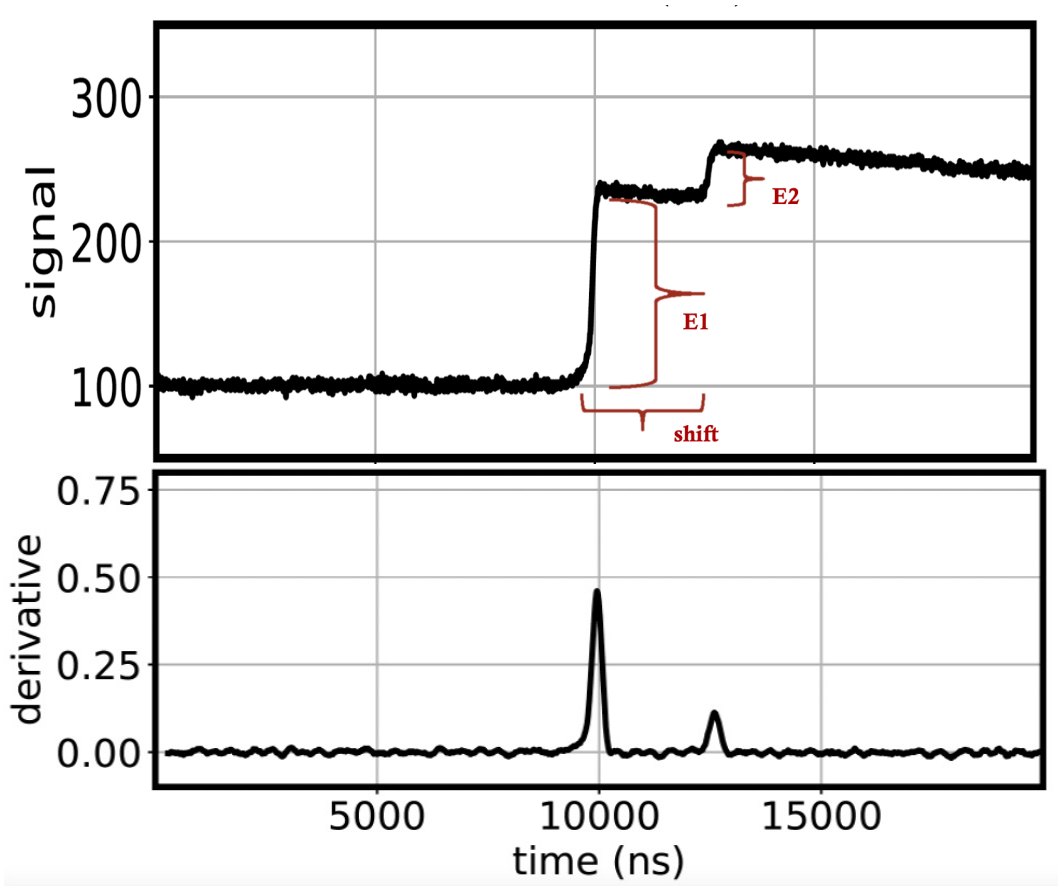


Animation credit: MJD collaboration

"LEGEND: Searching for 0νββ in 76-Ge"
By Brady Bos

Motivation for machine learning-based waveform study

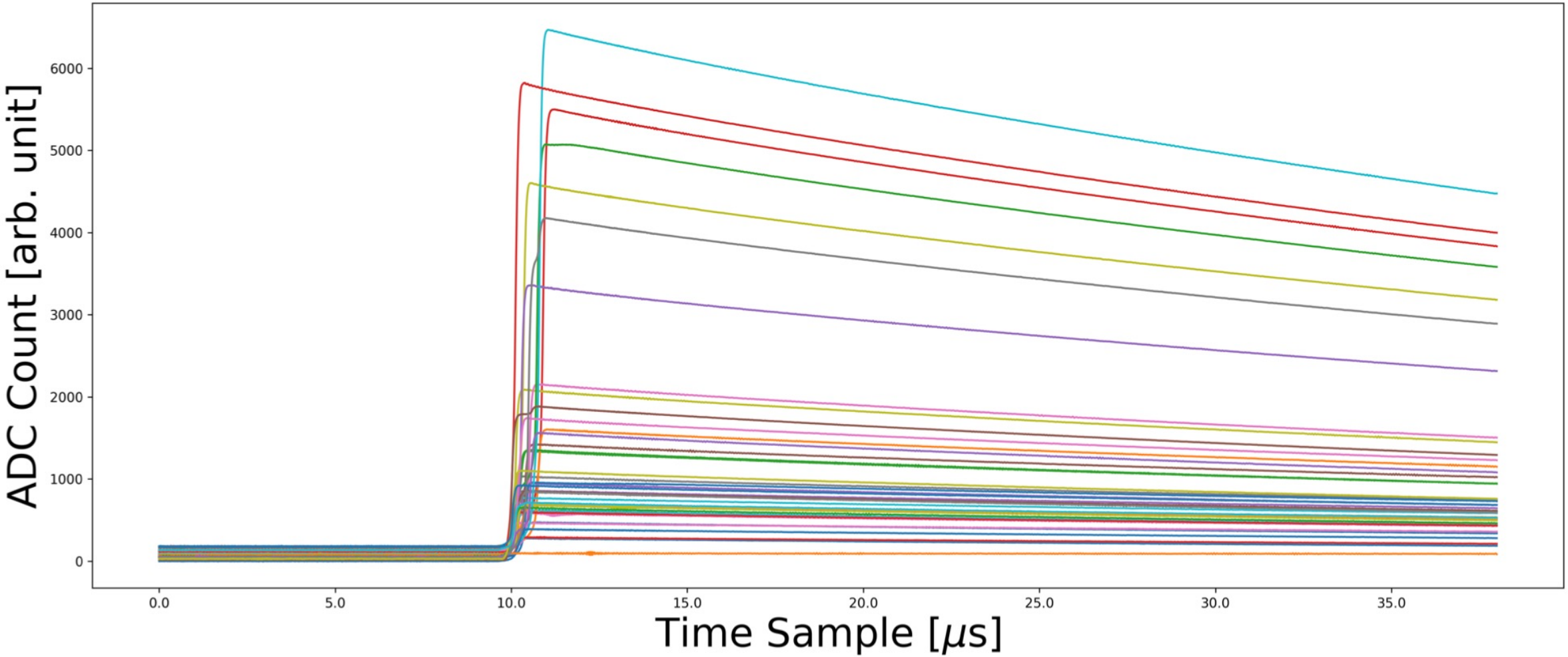




- In physics data, there is a small amount of expected pileup waveforms, e.g., the signature of isomeric gamma transition following cosmogenic production of certain isotopes
  - Look for shift and energy's peak ratio (E1/E2) for the signature.
  - This determination is implemented as regression in machine learning approach.

Data selection

❖ Used randomly selected 1% subset of DS6 ^{228}Th calibration runs of the MAJORANA DEMONSTRATOR detectors.

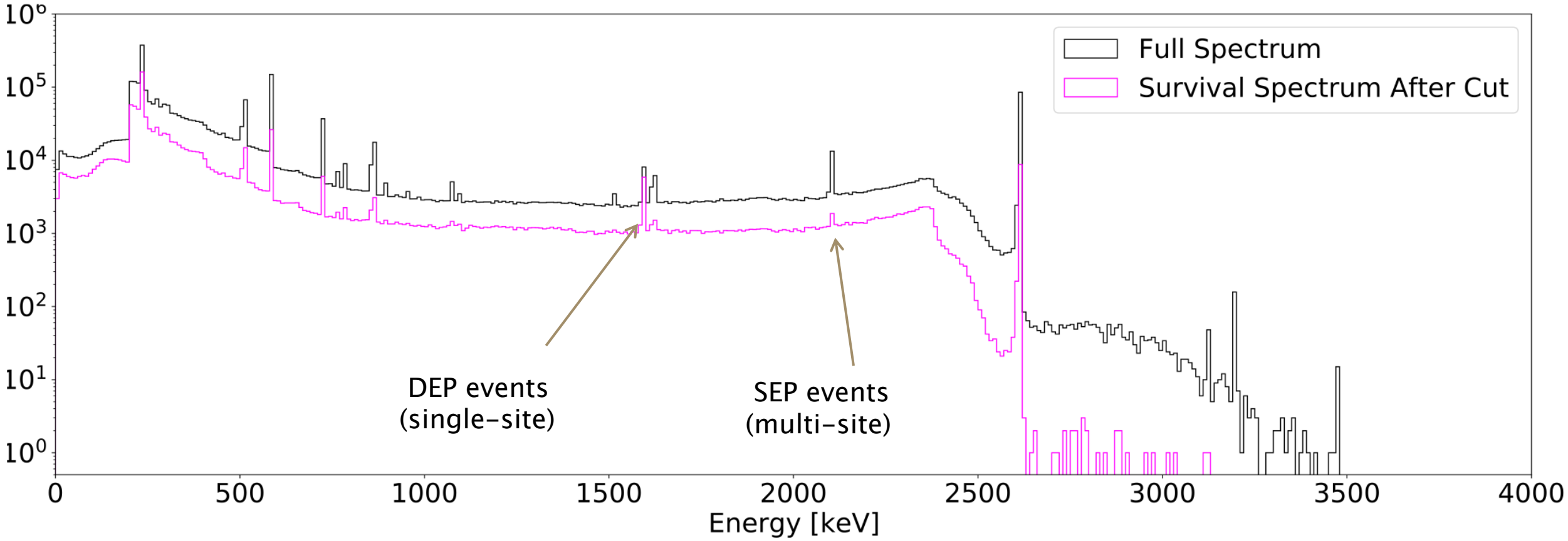


arXiv:2308.10856

Data selection

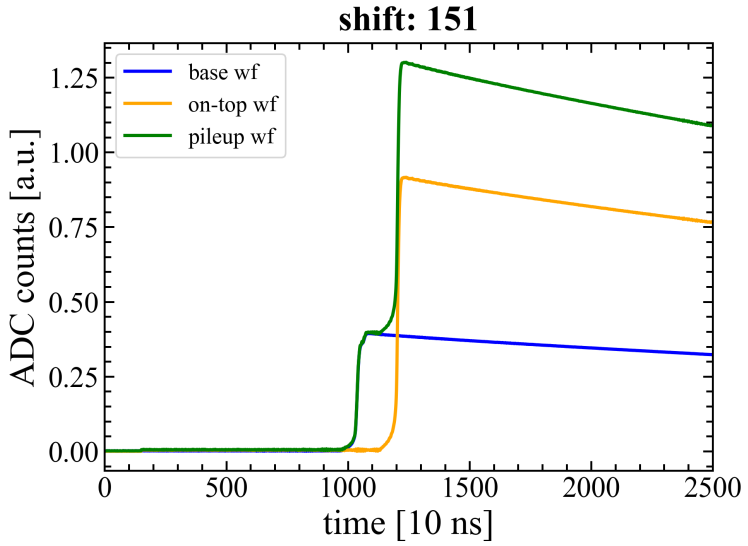
❖ Used randomly selected 1% subset of DS6 ^{228}Th calibration runs of the MAJORANA DEMONSTRATOR detectors.

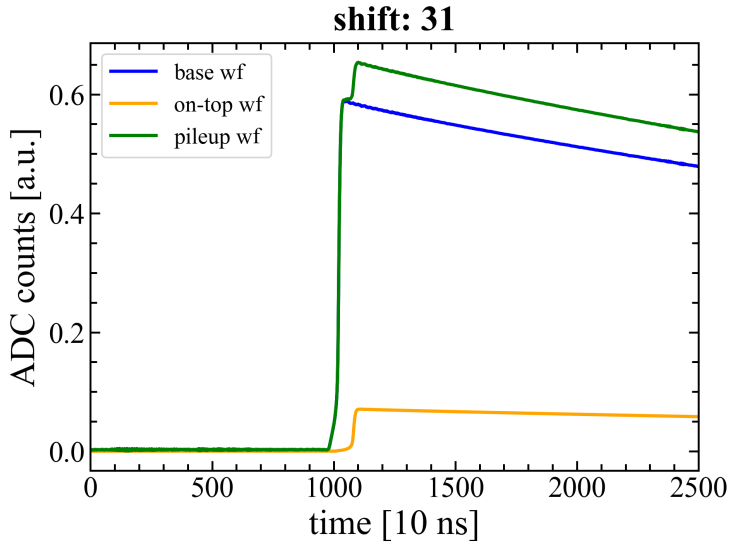
❖ Collect double escape peak (DEP) events based on energy cut

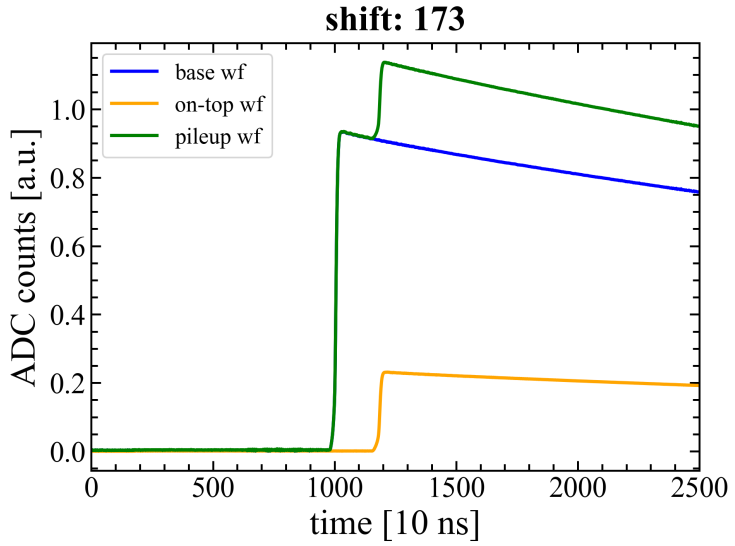


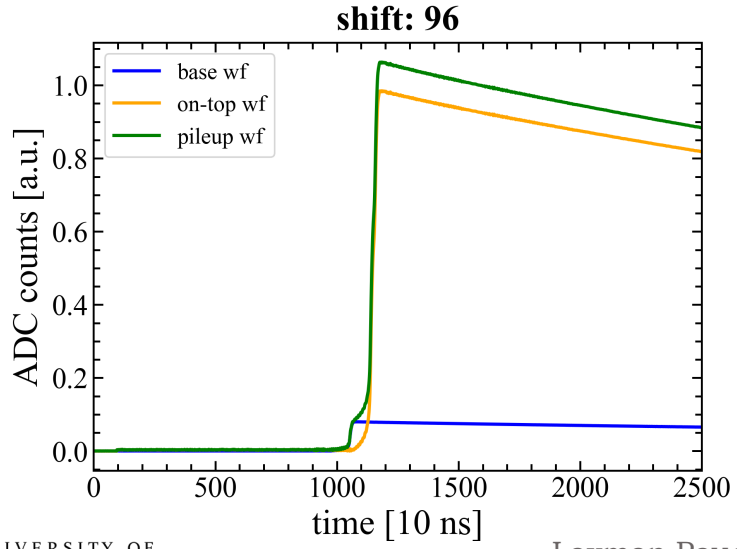
Simulated pileup waveforms dataset

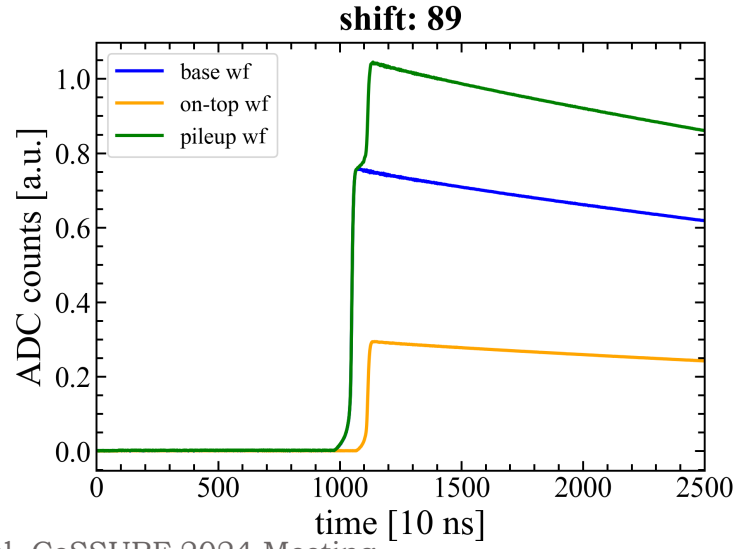
- Waveforms are single-site events (DEP) taken from the calibration data. Use these single waveforms to simulate pileup waveforms.

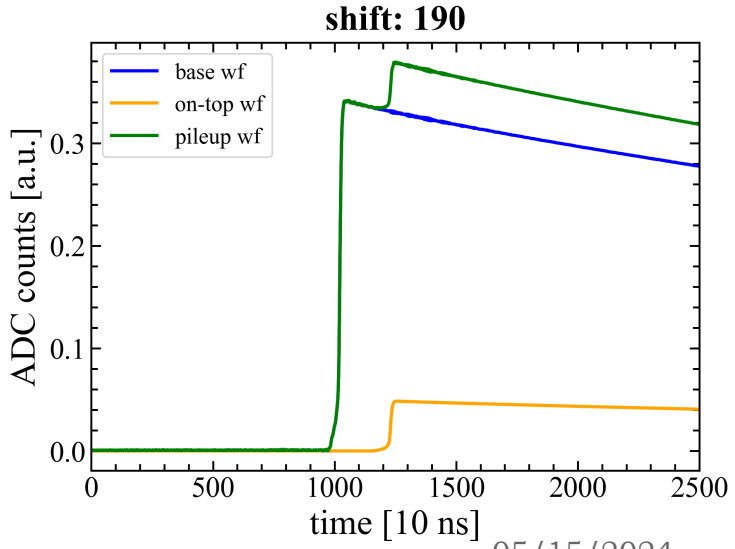






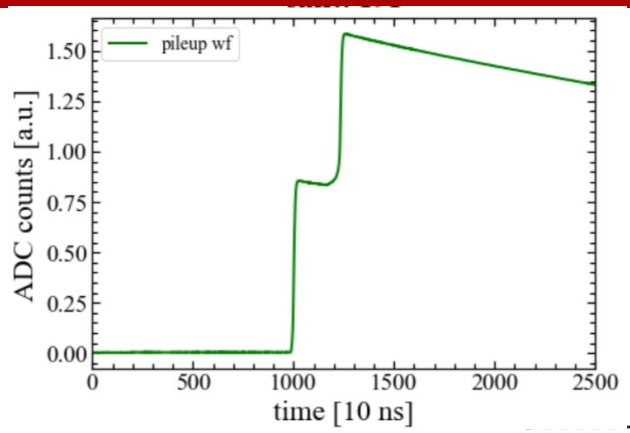


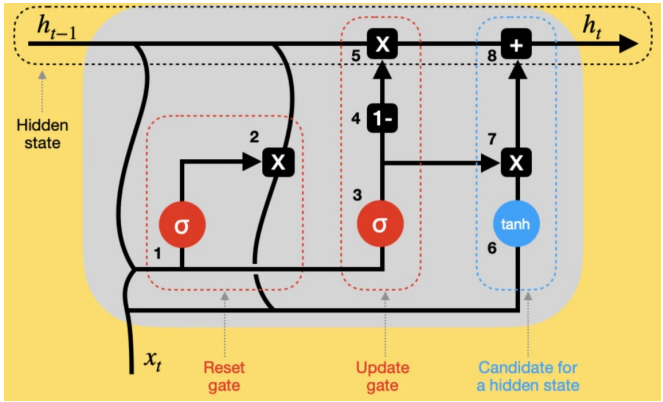




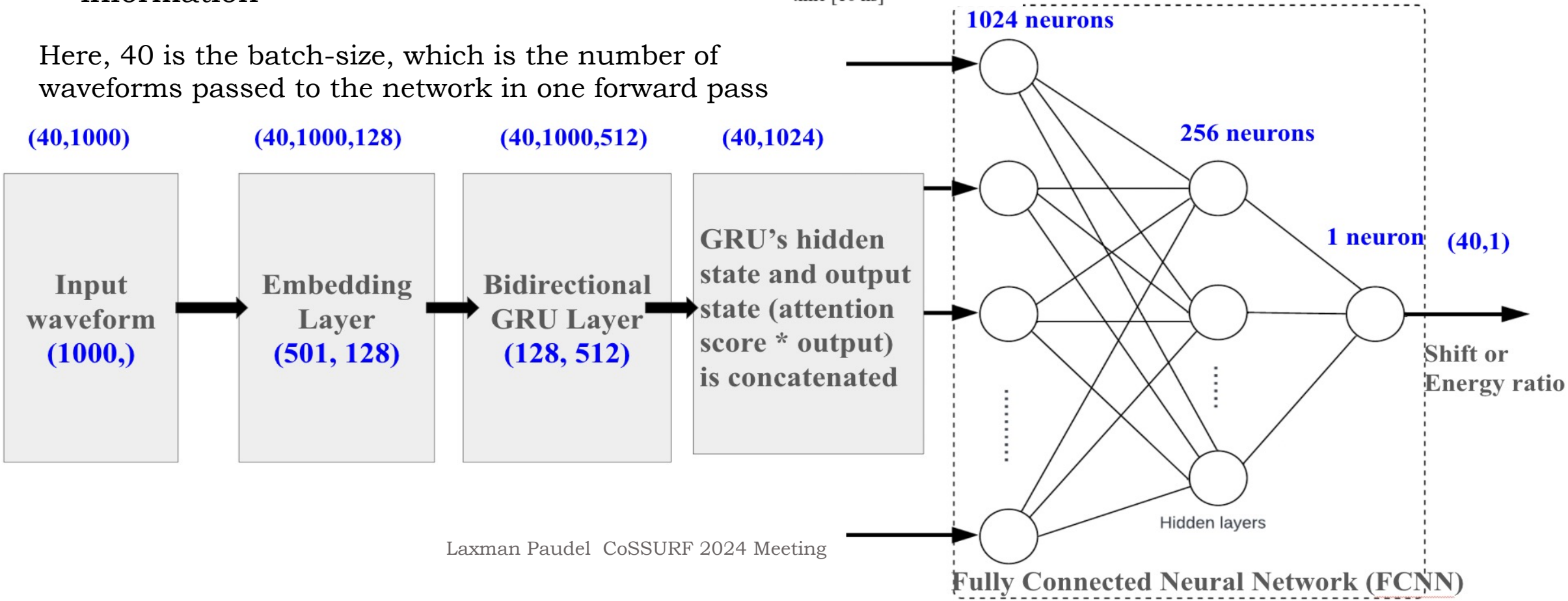
Recurrent Neural Network

- RNN is a canonical model for processing waveform data.
- Attention mechanism allows RNN to zoom in the part which contains most important information



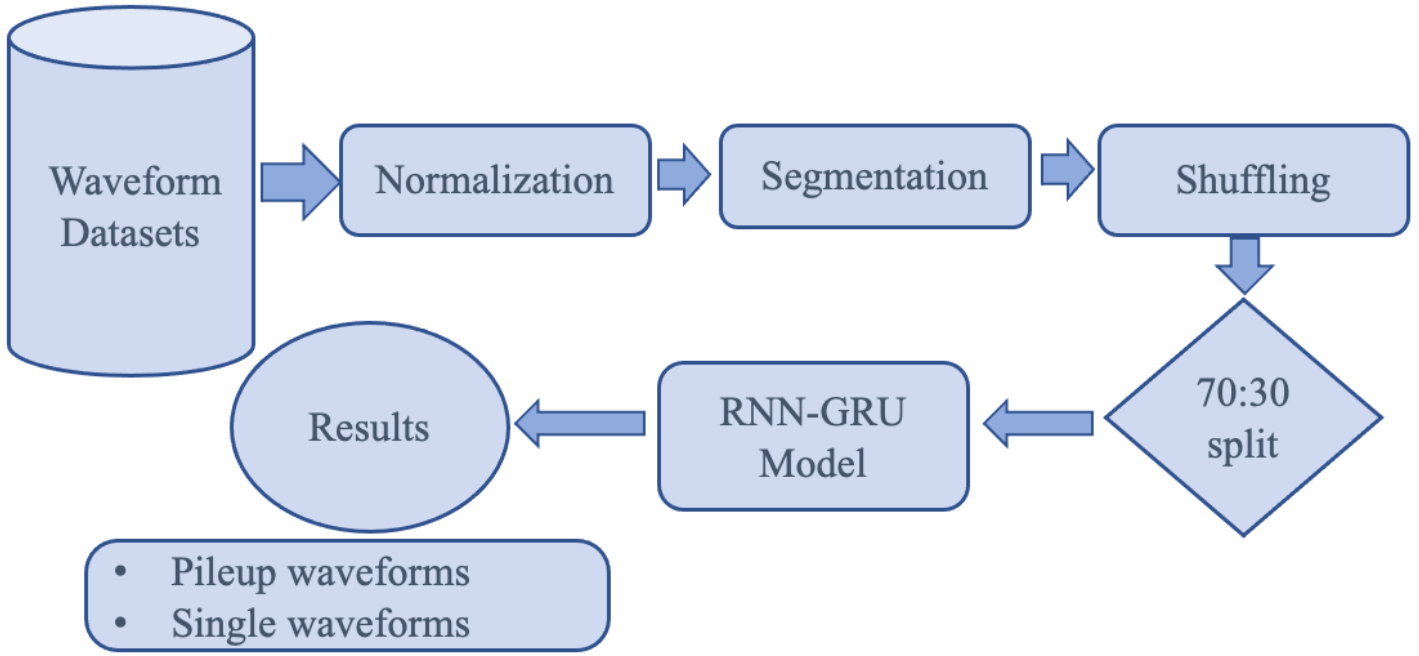


Here, 40 is the batch-size, which is the number of waveforms passed to the network in one forward pass



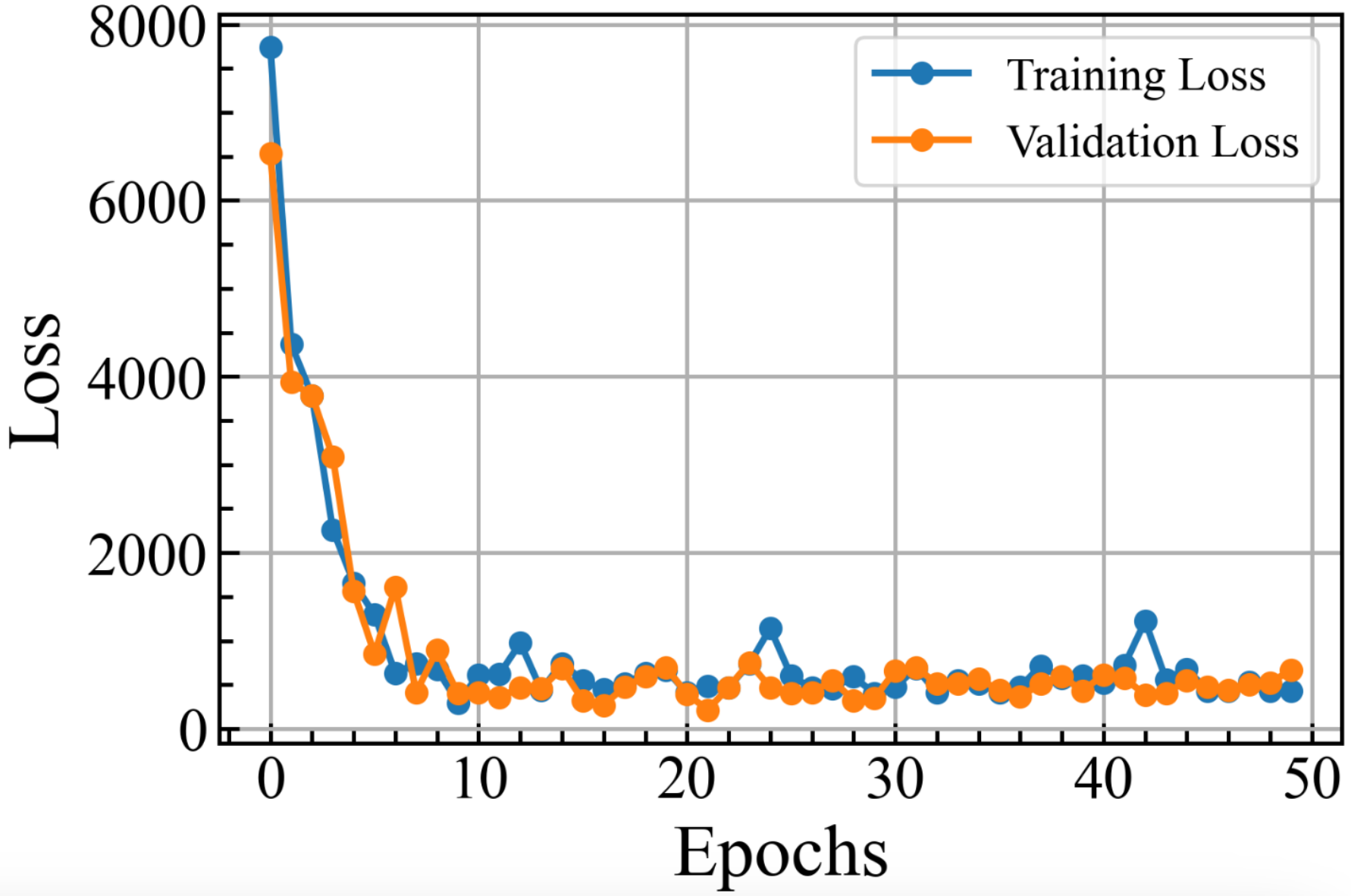
RNN training for classification

- Waveforms are normalized
- Training data:
  - 70,000 waveforms are used for training
  - For classification: Labeling of waveform (single waveform: label = 0, pileup waveform: label = 1)



- Testing data:
  - 30,000 waveforms are used for testing
  - For classification: value close to 1 or 0 means the network thinks the input waveform is pileup or single waveforms

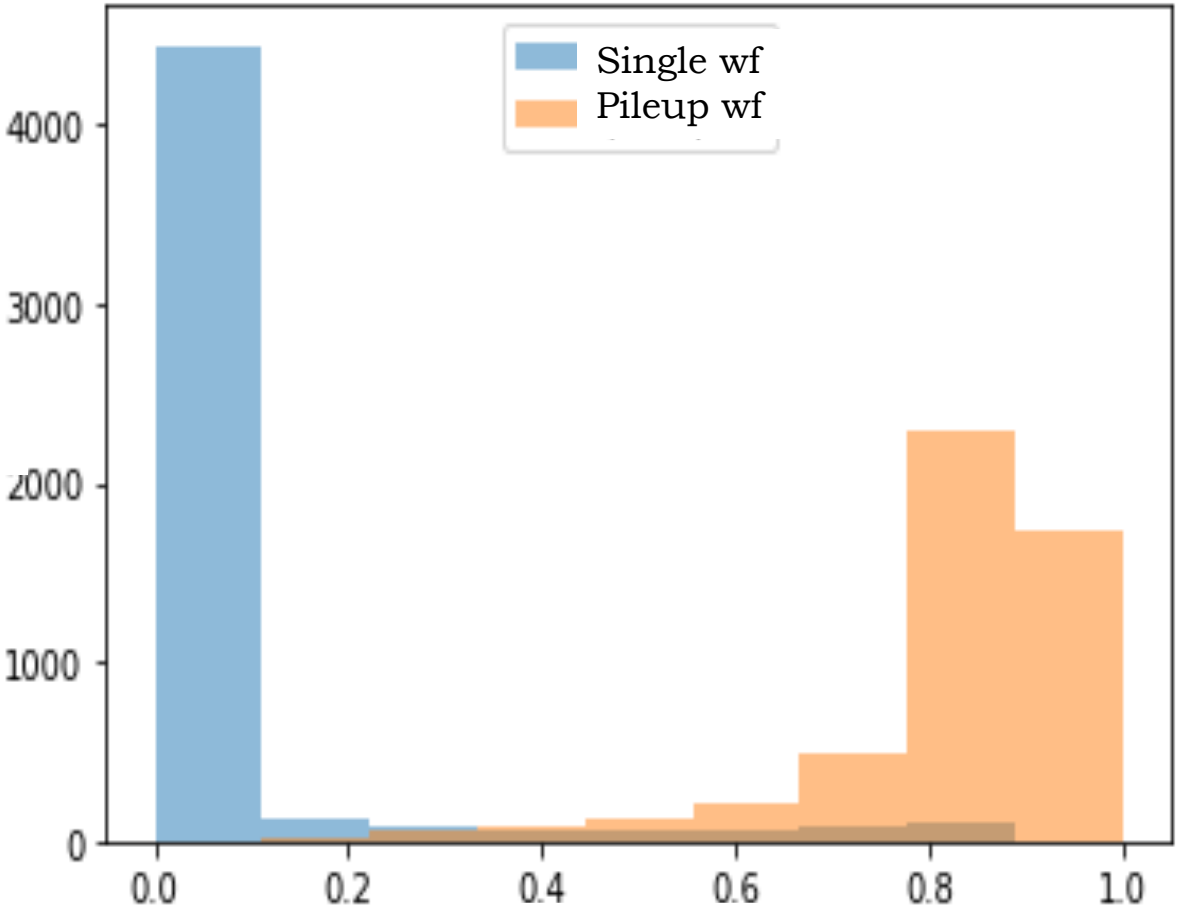
Learning curve



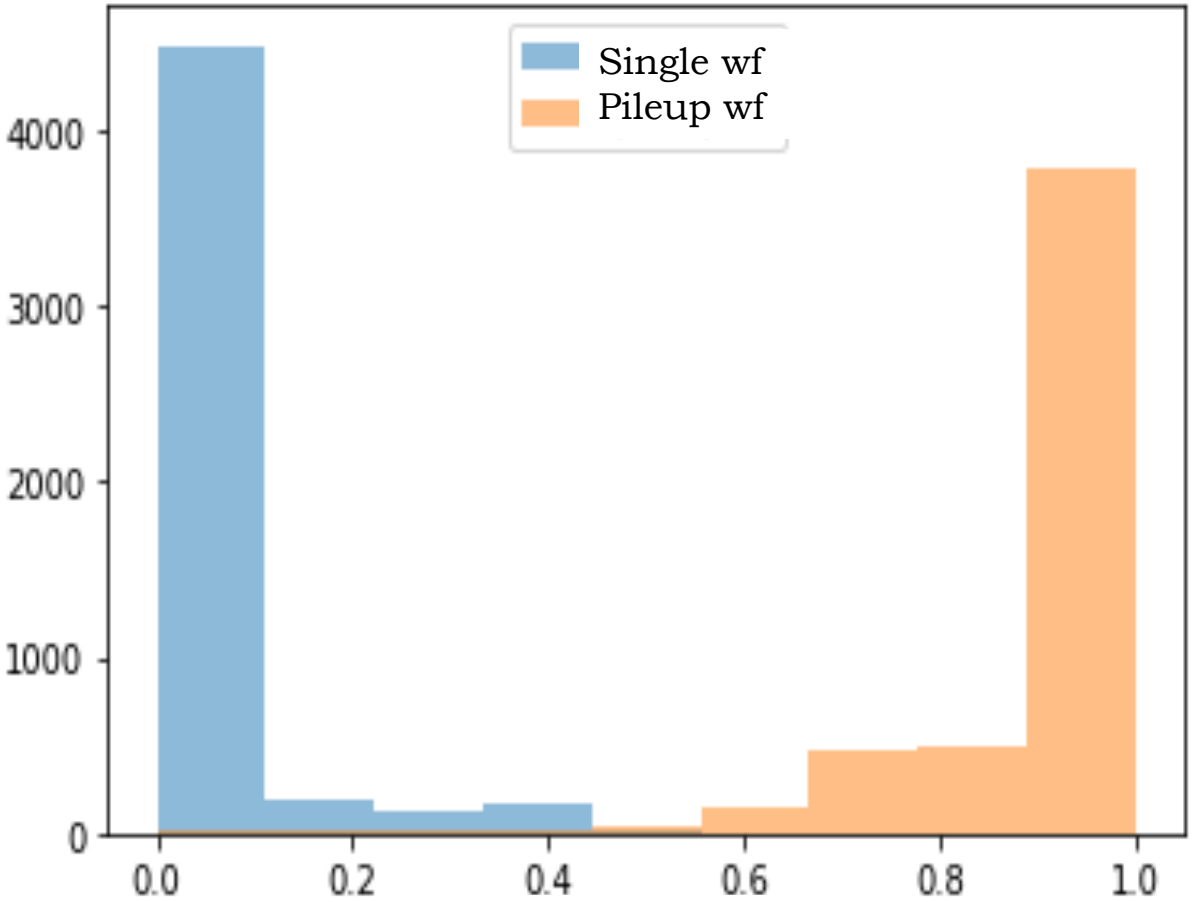
RNN output for pileup and single waveforms

- RNN outputs for a maximum shift range up to 50 ns and a maximum shift range up to 2000 ns.

Shift value: 50 ns



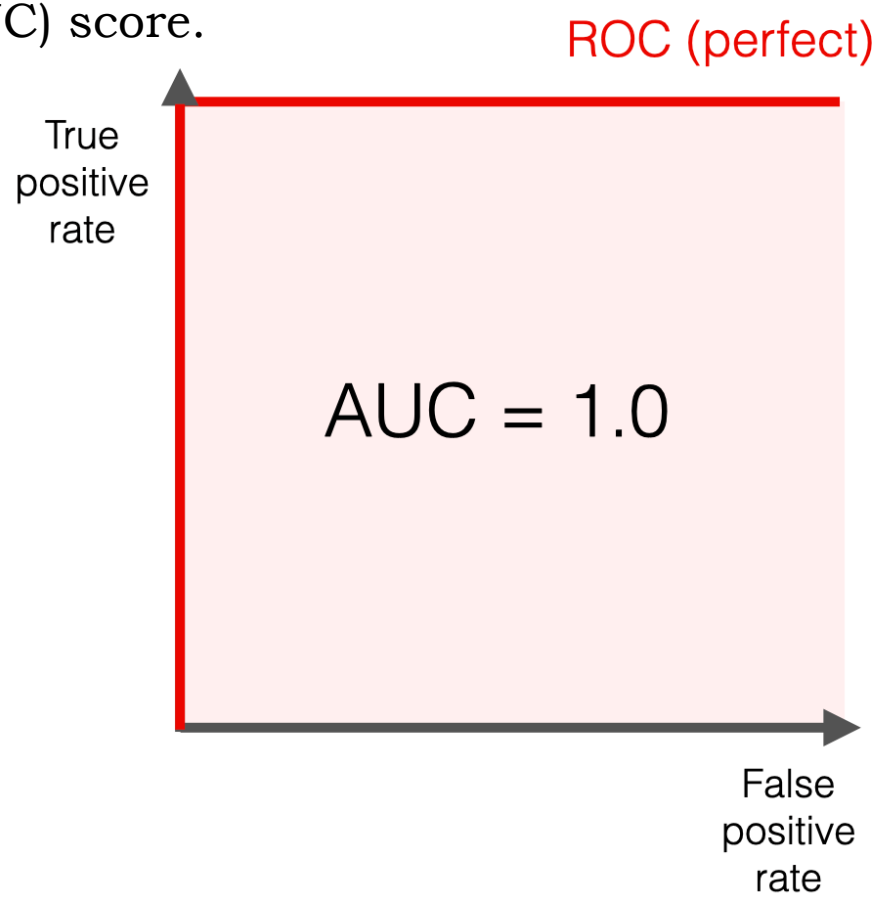
Shift value: 2000 ns

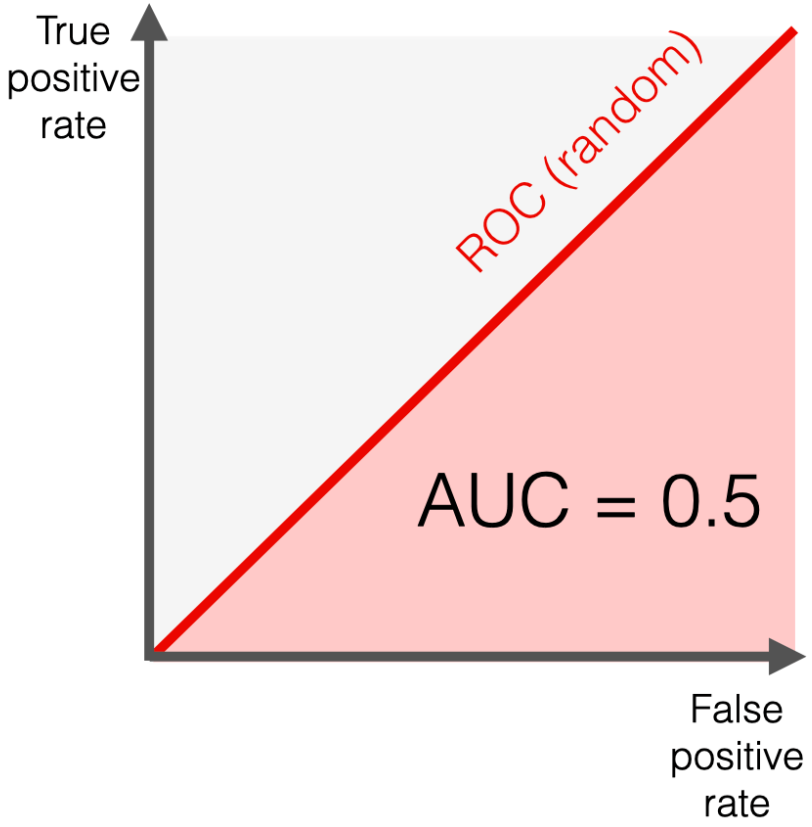


Result: classification between single and pileup waveform

- Receiver Operating Characteristic curve (ROC) is a plot between True Positive rates (TPR) and the False Positive rates (FPR).

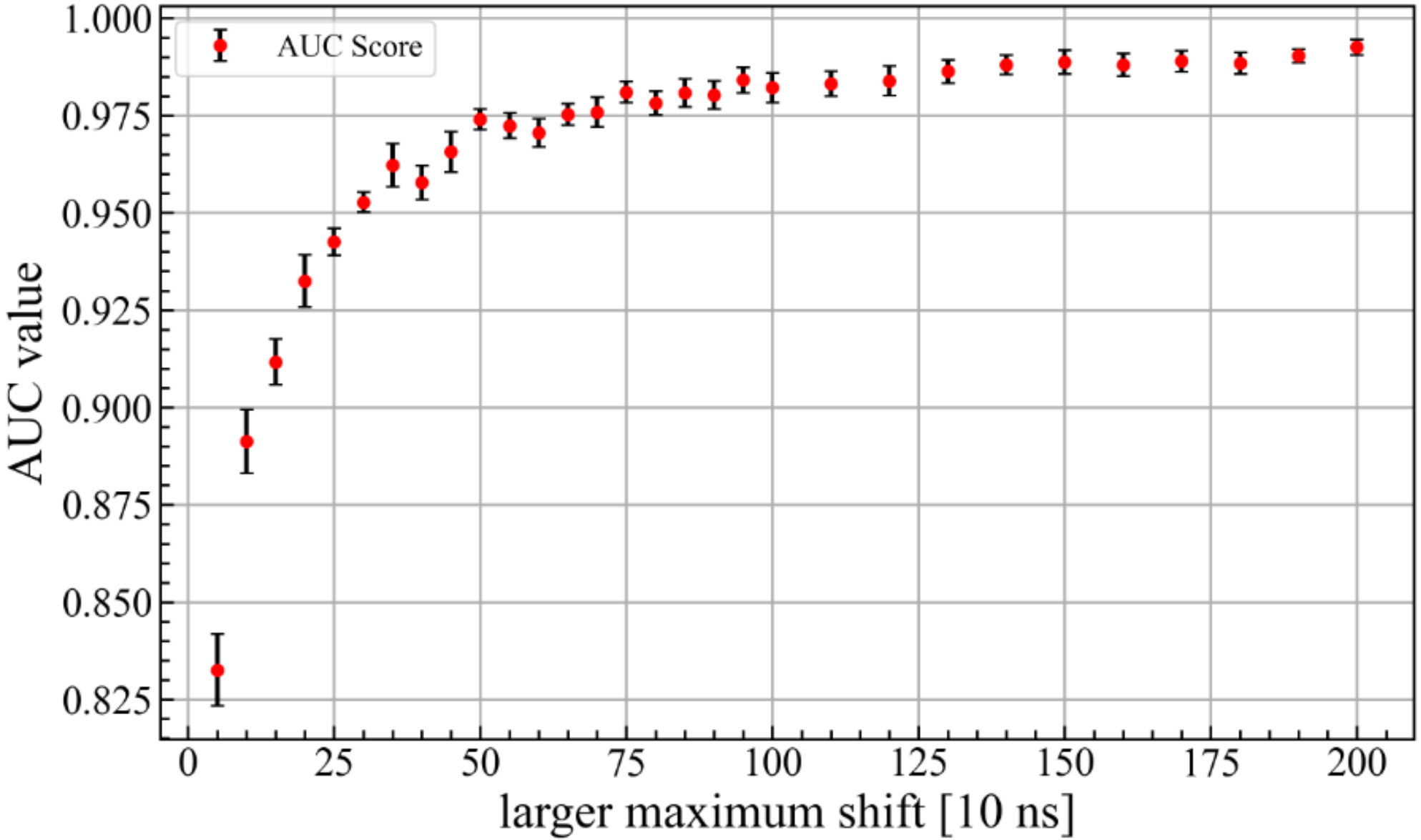
- The performance of classification power of the model can be tested using Area Under the ROC Curve (AUC) score.



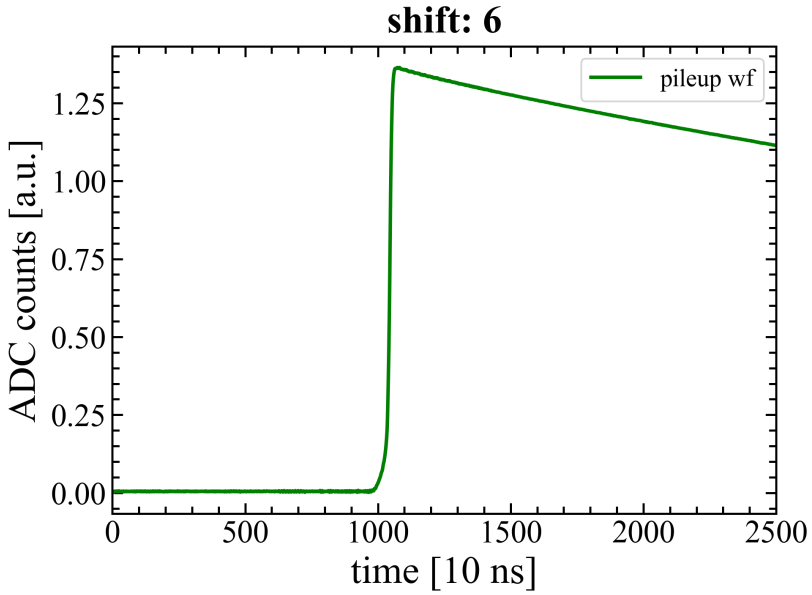


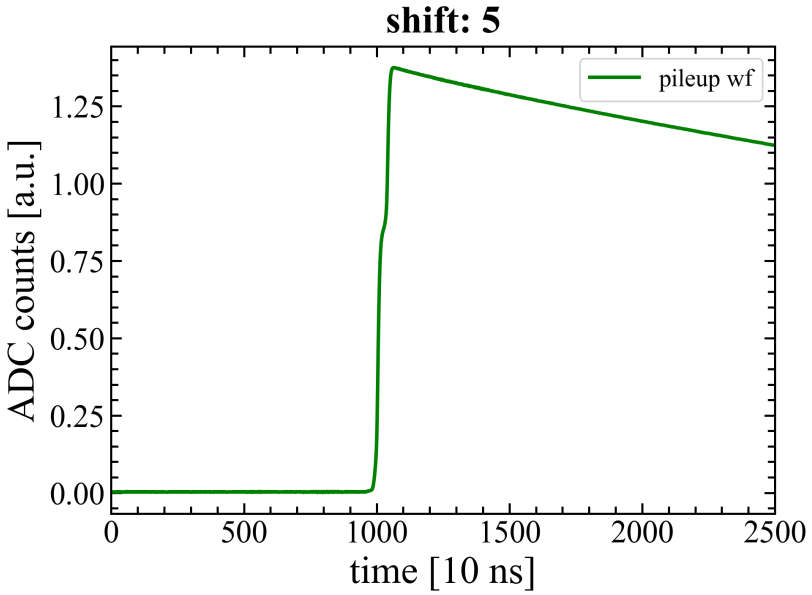
https://www.evidentlyai.com/classification-metrics/explain-roc-curve

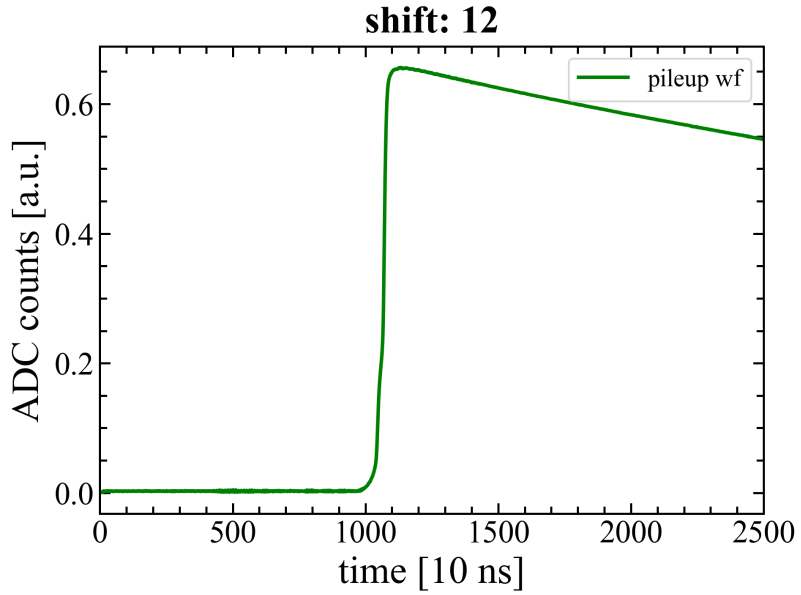
Result: classification between single and pileup waveform

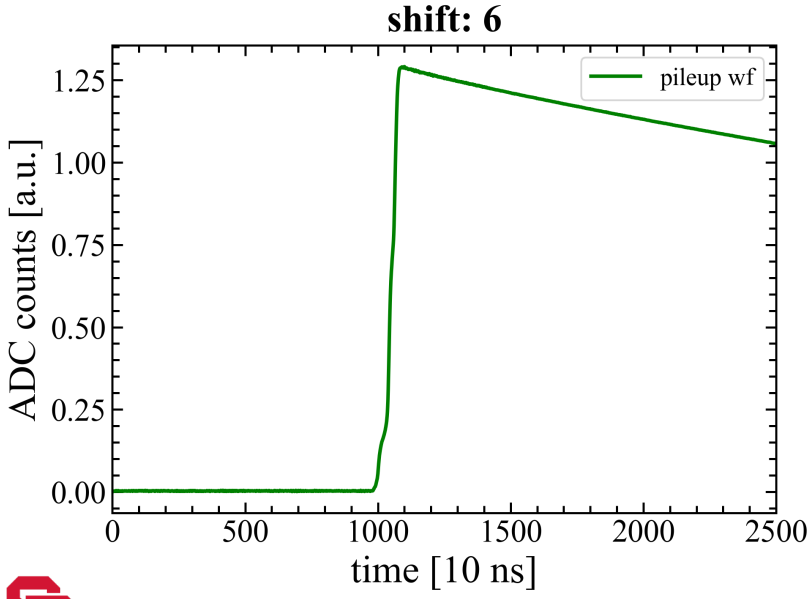


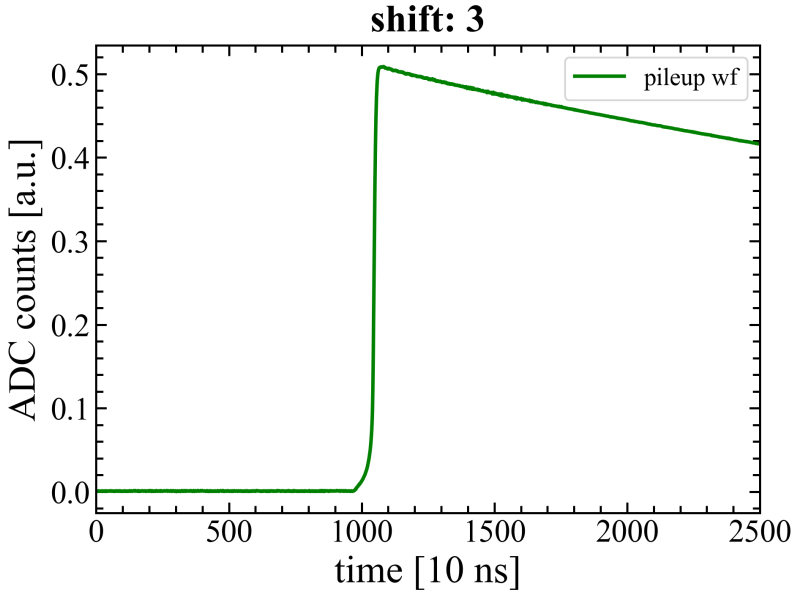
Result: classification between single and pileup waveform

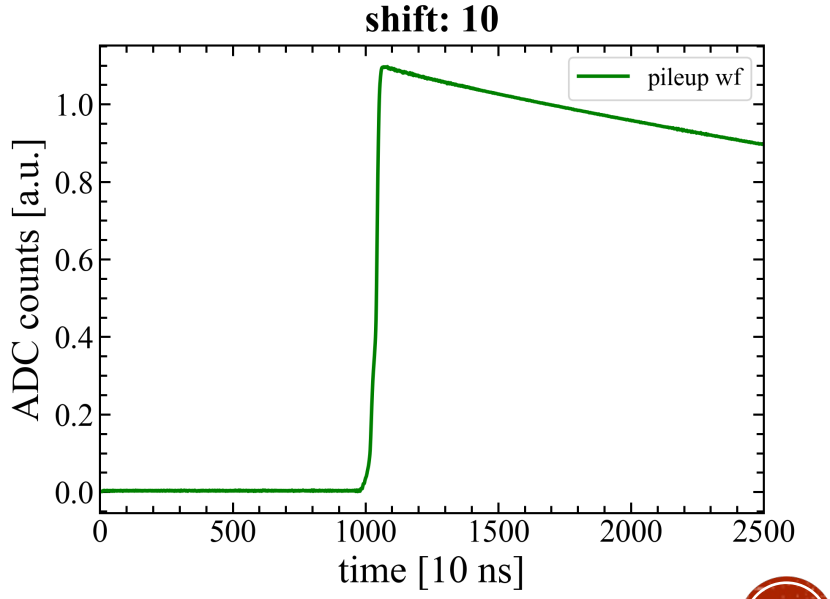








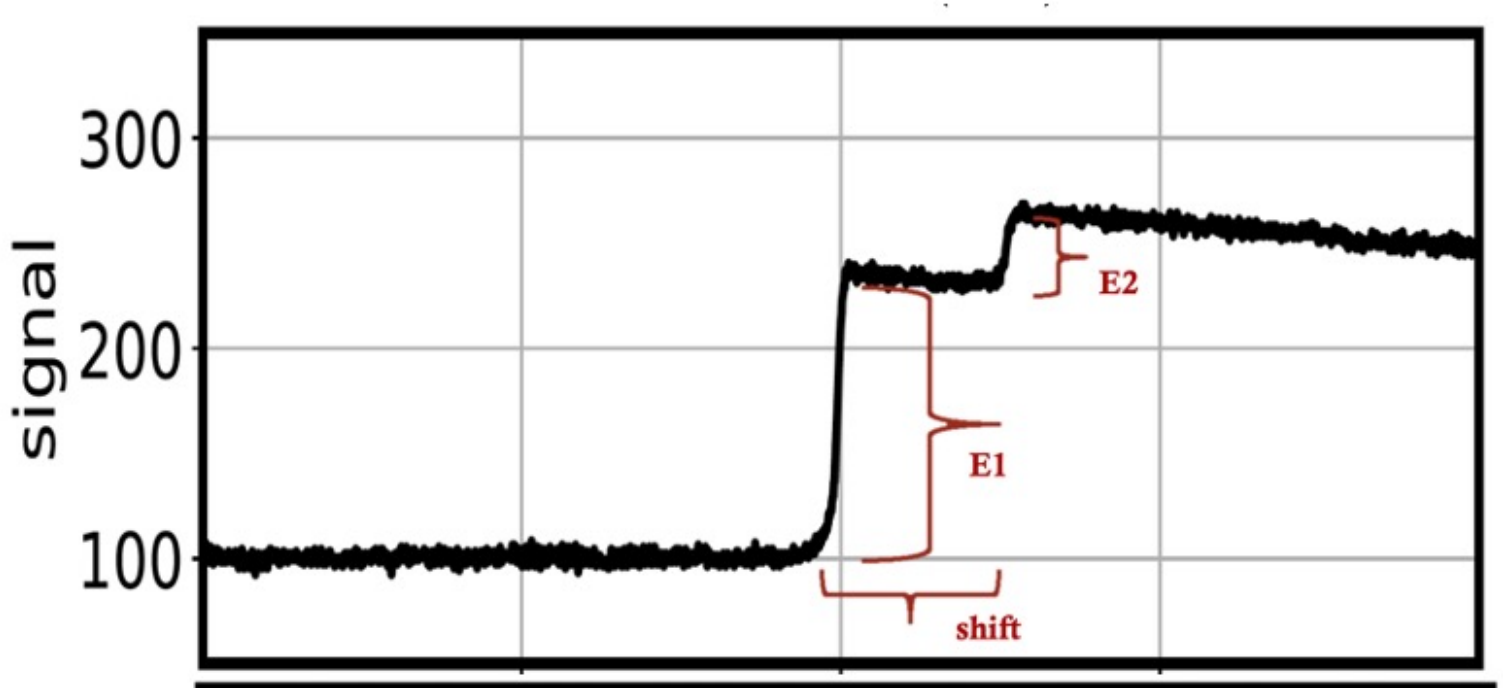




Recap and a new question

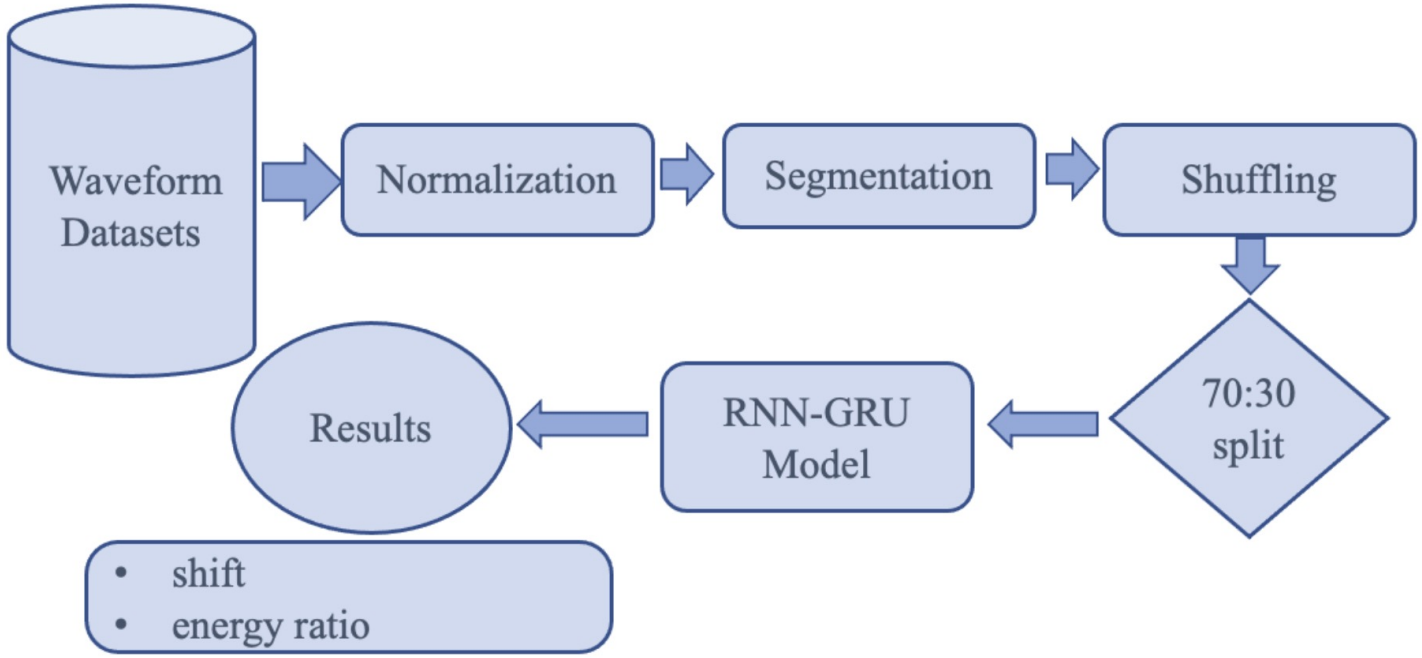
- The network has high efficiency to distinguish pileup waveforms from single waveforms

- Can the network also tell how much time shift is between the 2 steps in a pileup?
  - ML-based regression



RNN training for regression

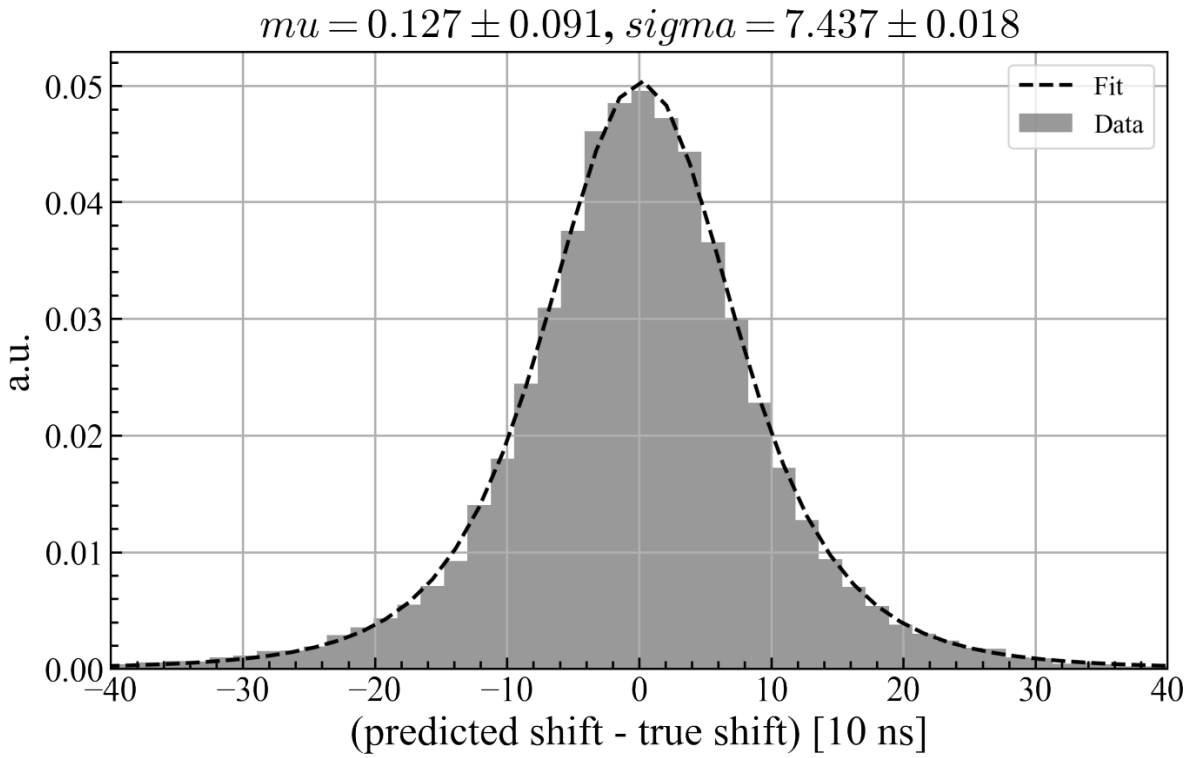
- Waveforms are normalized
- Training data:
  - 70,000 waveforms are used for training
  - For regression: we provide the values for shift and energy ratio of pileup waveforms

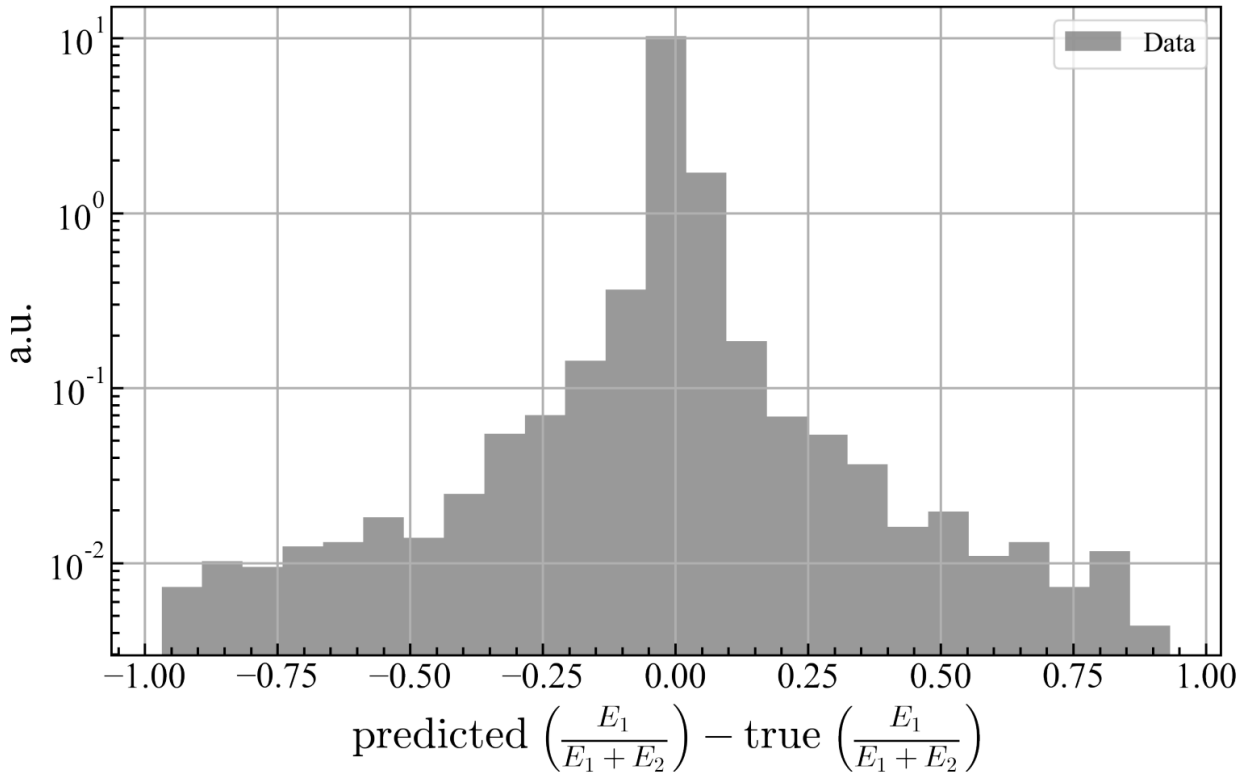


- Testing data:
  - 30,000 waveforms are used for testing
  - For regression: network gives the values for shift and energy ratio

Result: parameter regression in RNN

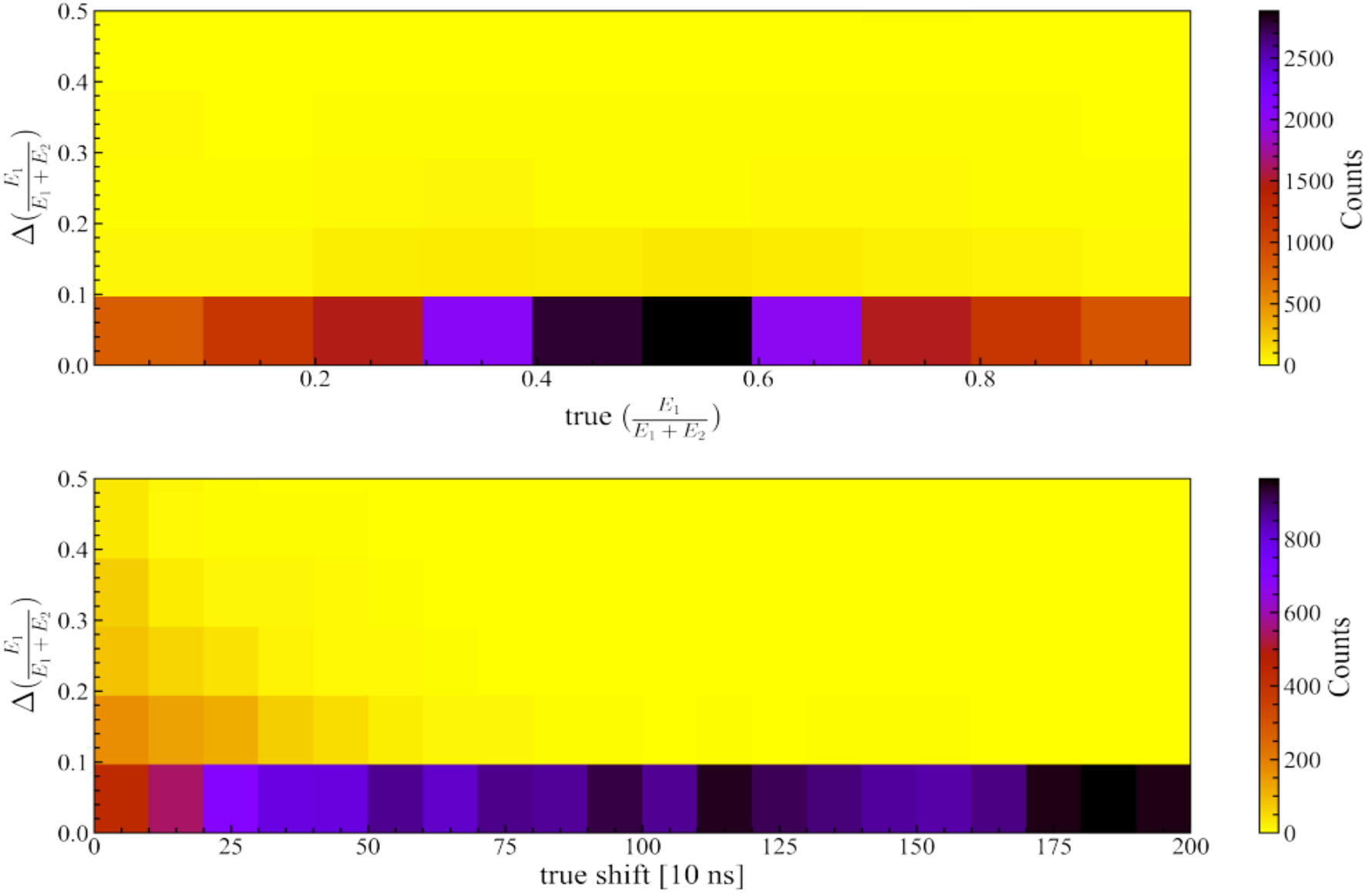
- Difference between true shift and predicted shift, as well as difference between true energy ratio and predicted energy ratios are studied from the regression.





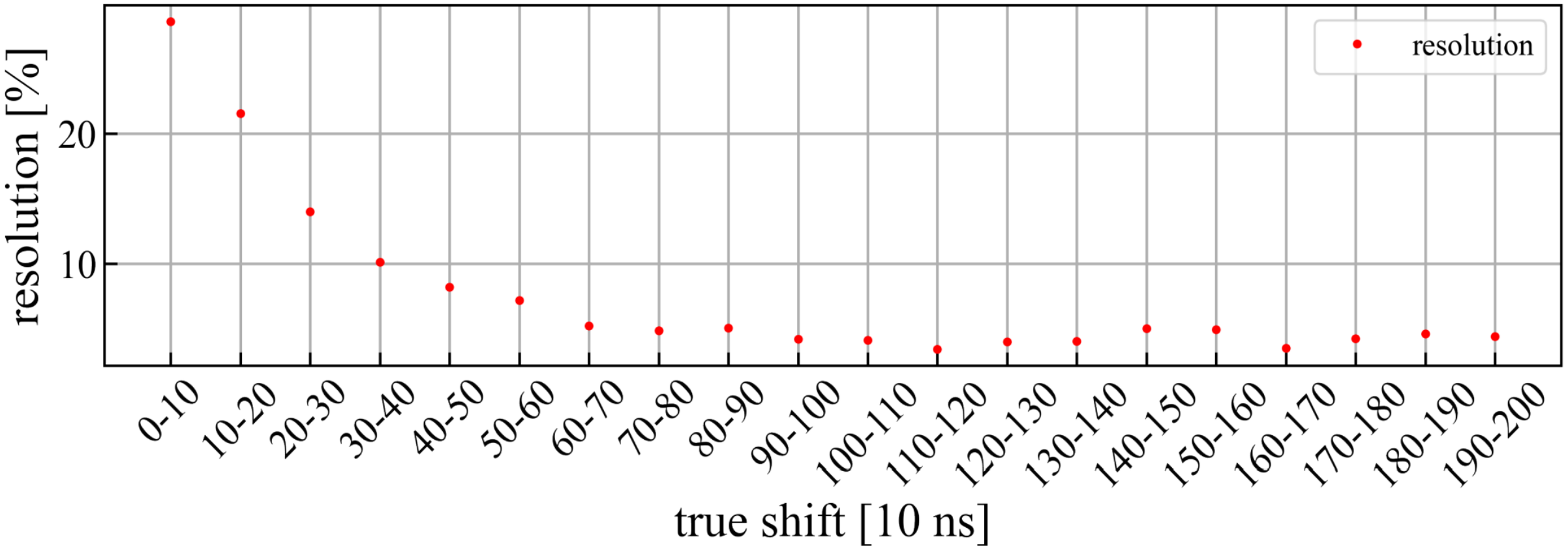
- Good parameter determination performance.

Result: parameter regression in RNN

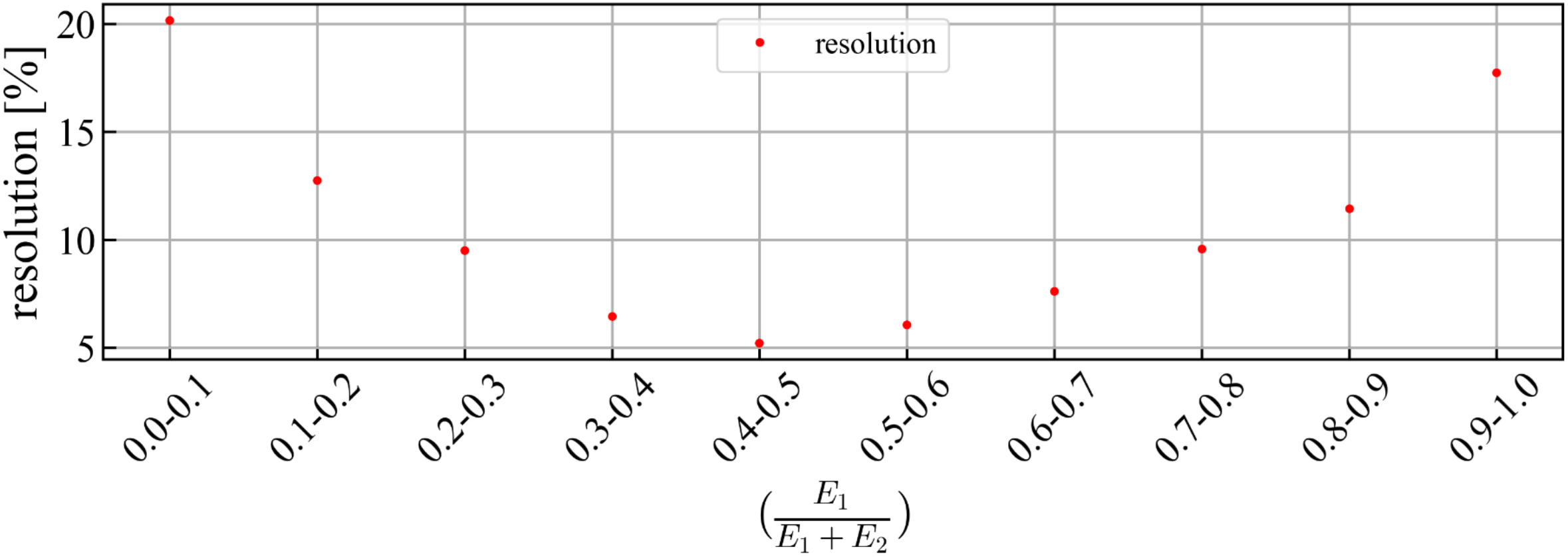


Result: parameter regression in RNN

- We define energy ratio resolution as the percentage standard deviation of difference between true and predicted energy ratio across the test dataset.
- The energy ratio resolution is improved with increasing time shifts, as expected.



Result: parameter regression in RNN



- This agrees with our expectation: near 0.5 energy ratio, the 2-sites of pileup events is the most obvious; when the energy ratio becomes too large (0.9) or too small (0.1), pileup events become similar to single site events making energy reconstruction more difficult.

Summary and outlook:

- In single waveform and pileup binary classification, RNN shows a good performance, particularly better at larger relative shifts range than smaller relative shifts range
  - Extremely small shifts are very similar to single-site vs. multiple site classification.

- Presented results are for simulated waveforms. We can apply the trained networks in real dataset, for example, energy calibration data to look for random pile-up events.

- We can apply the model to look for the signature of isomeric gamma transition in data.

"Pushing Rare Event Search to the Limit with Machine Learning Algorithms"
Plenary talk by Dr. Aobo Li

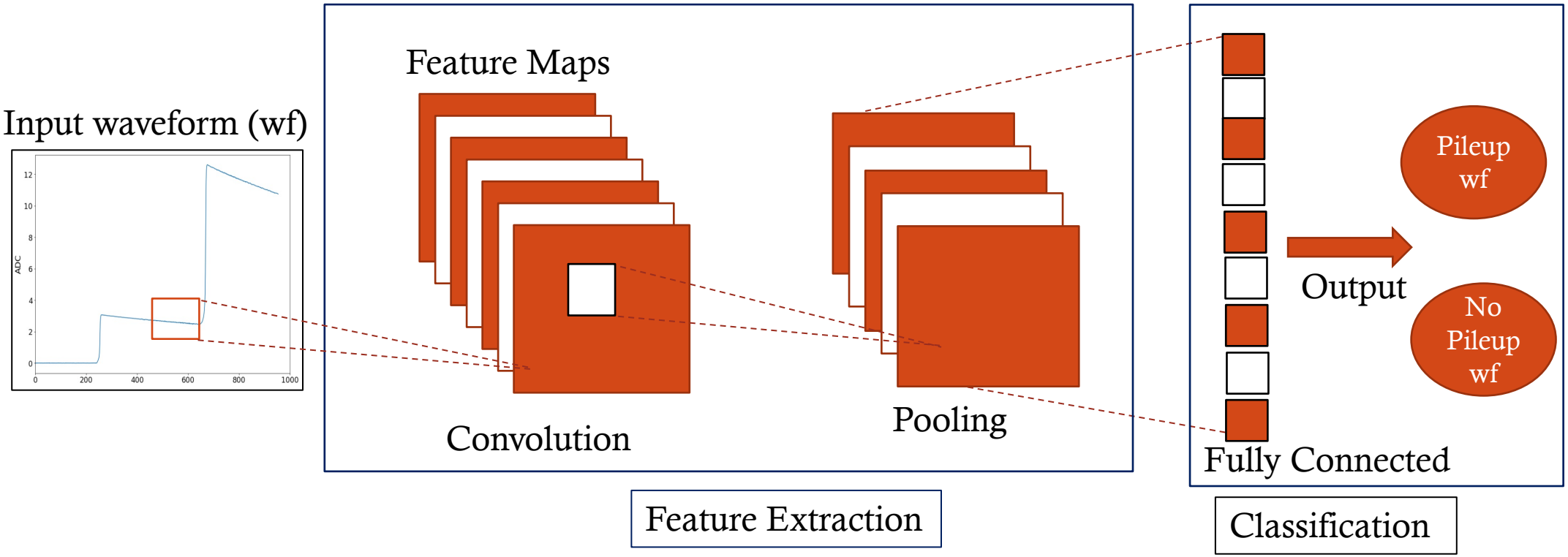
"Majorana Demonstrator Data Release for AI/ML Applications"
arXiv:2308.10856
by MJD Collaboration

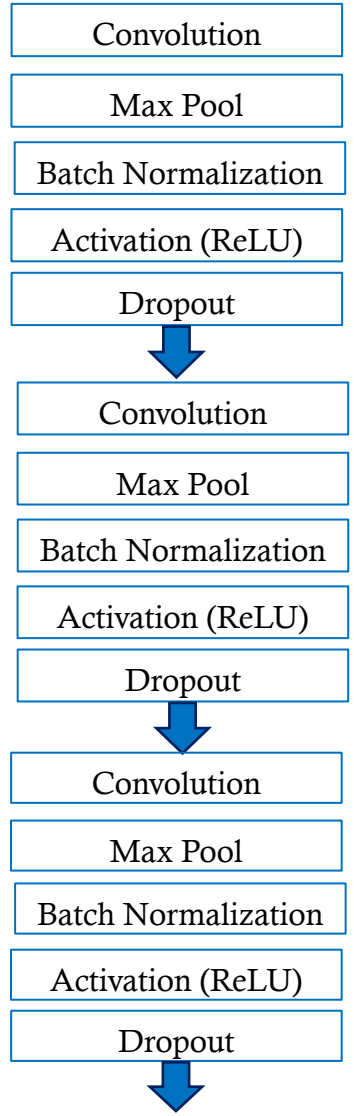
Acknowledgements

- Dr. Aobo Li
- MAJORANA DEMONSTRATOR collaboration
- LEGEND collaboration

Backup

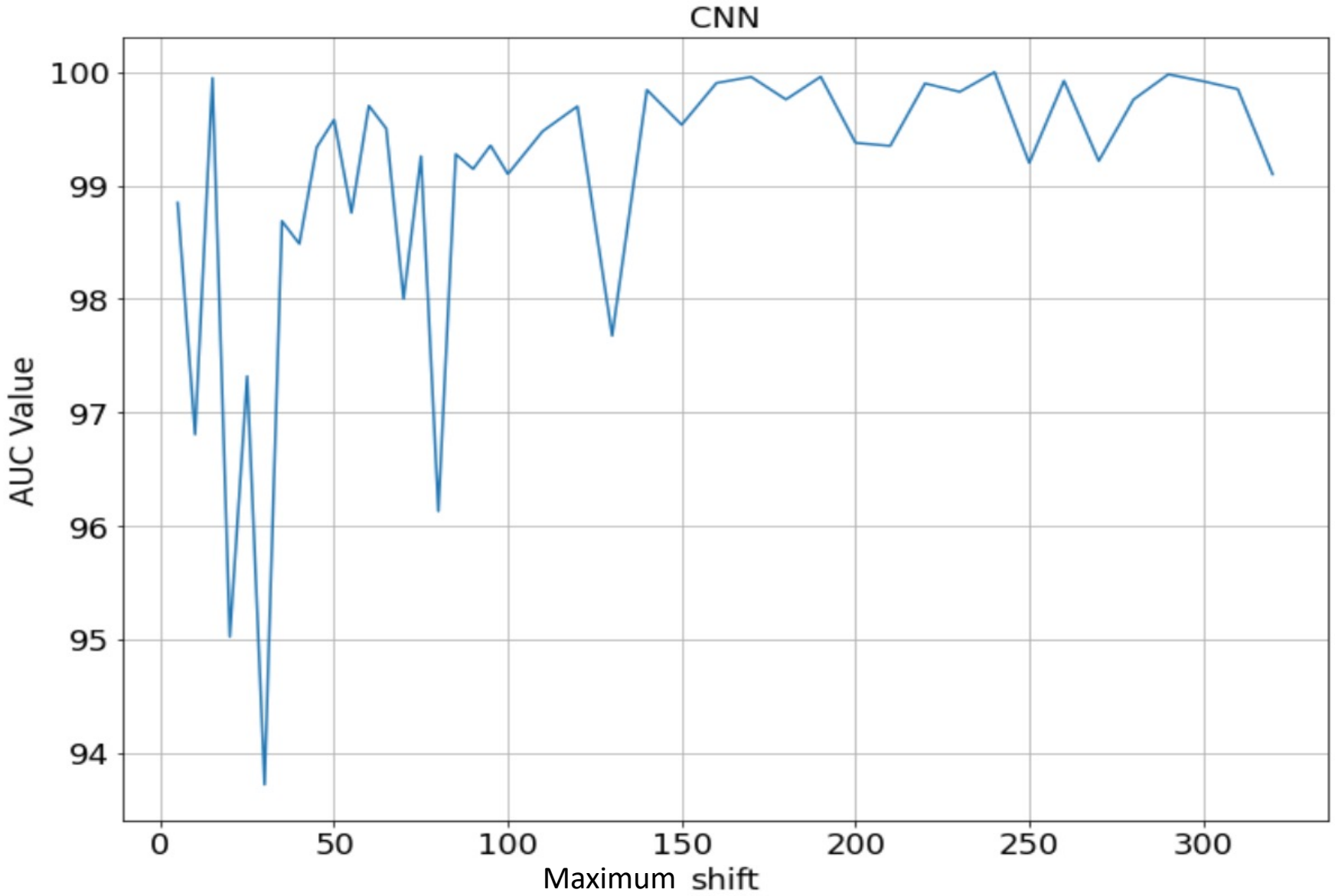
Convolutional Neural Network (CNN)





- Each convolutional layer contains a set of kernels, or filters, or feature detector.
- Convolution is the same as the filtering process in Ge detector, except that the shape of filter changes with respect to data.
- If we stack another convolutional layers on the previous layer, then we can increase the reception field of CNN
- CNN is better for image classification.

Classification performance of CNN



- Good performance for large maximum shifts
- Unstable performance for small maximum shifts