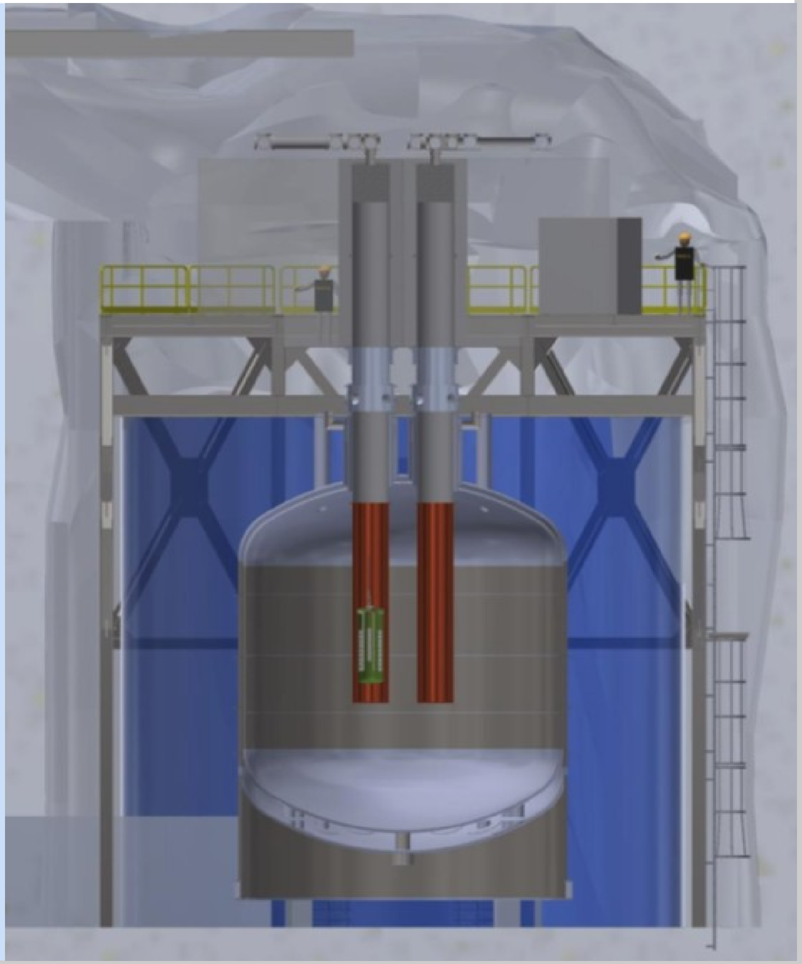
# Faithful Pulse Shape Analysis for Germanium Detectors using Feature Importance Supervision

LEGEND

Large Enriched
Germanium Experiment
for Neutrinoless ββ Decay



Katharina Kilgus

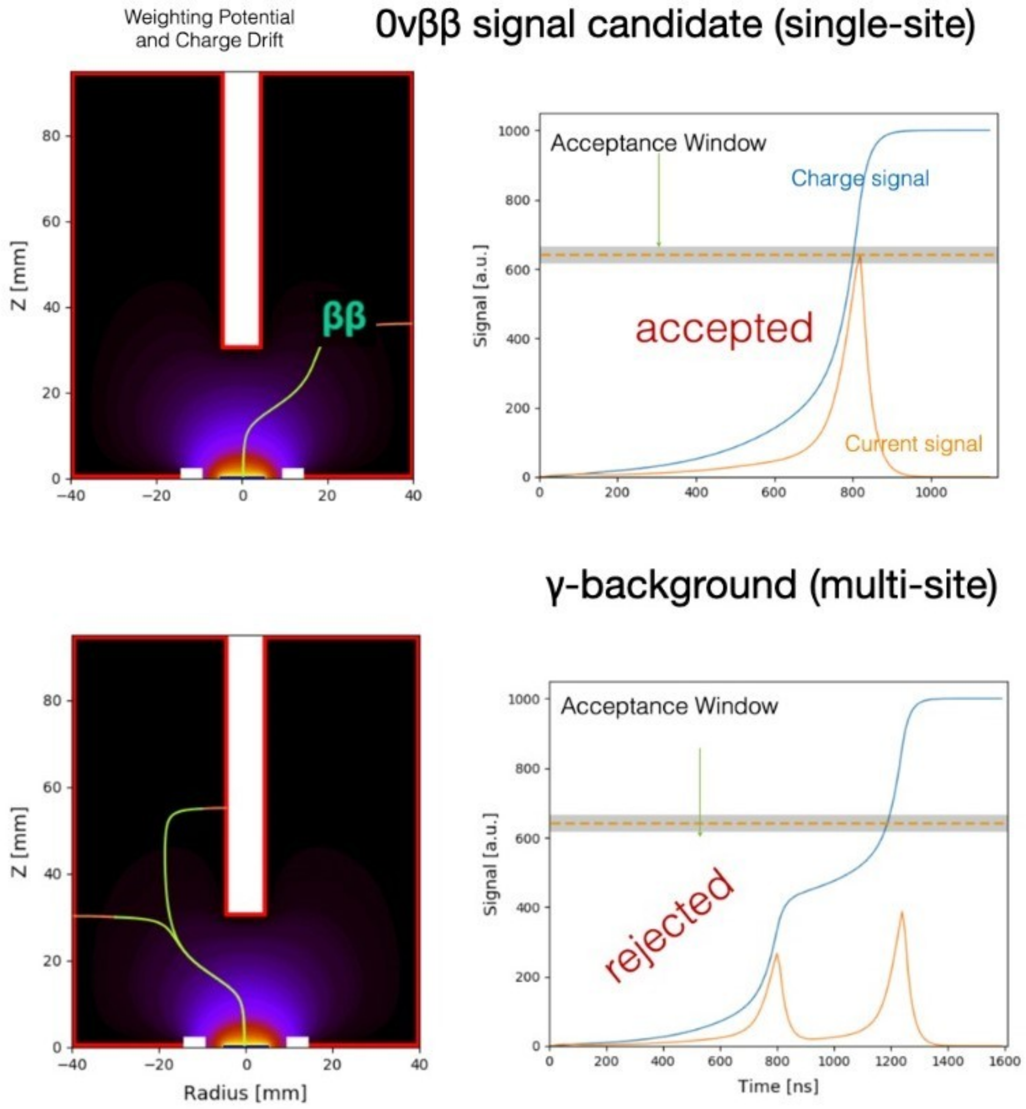
08/25/23

NPML 2023, Boston

# Outline

- Short about HPGe Detectors
- Pulse Shape Analysis (PSA)
- Feature Importance Supervision (FIS)
- Application to Germanium Signals
- Further possibilities by FIS

# Short about HPGe Detectors



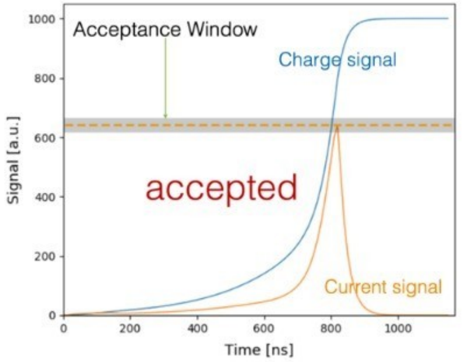

- Good requirements for Pulse Shape Analysis
- Classic analysis is done by A/E

**Main Task to keep in Mind:**

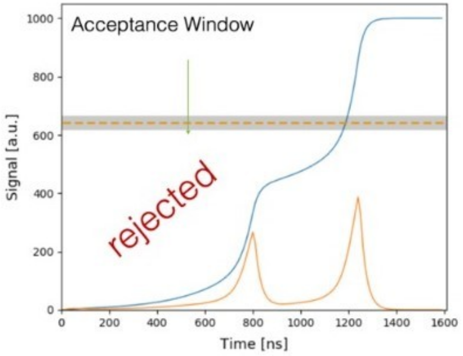
Discriminate MSE (one background-type) from SSE (signal-like)

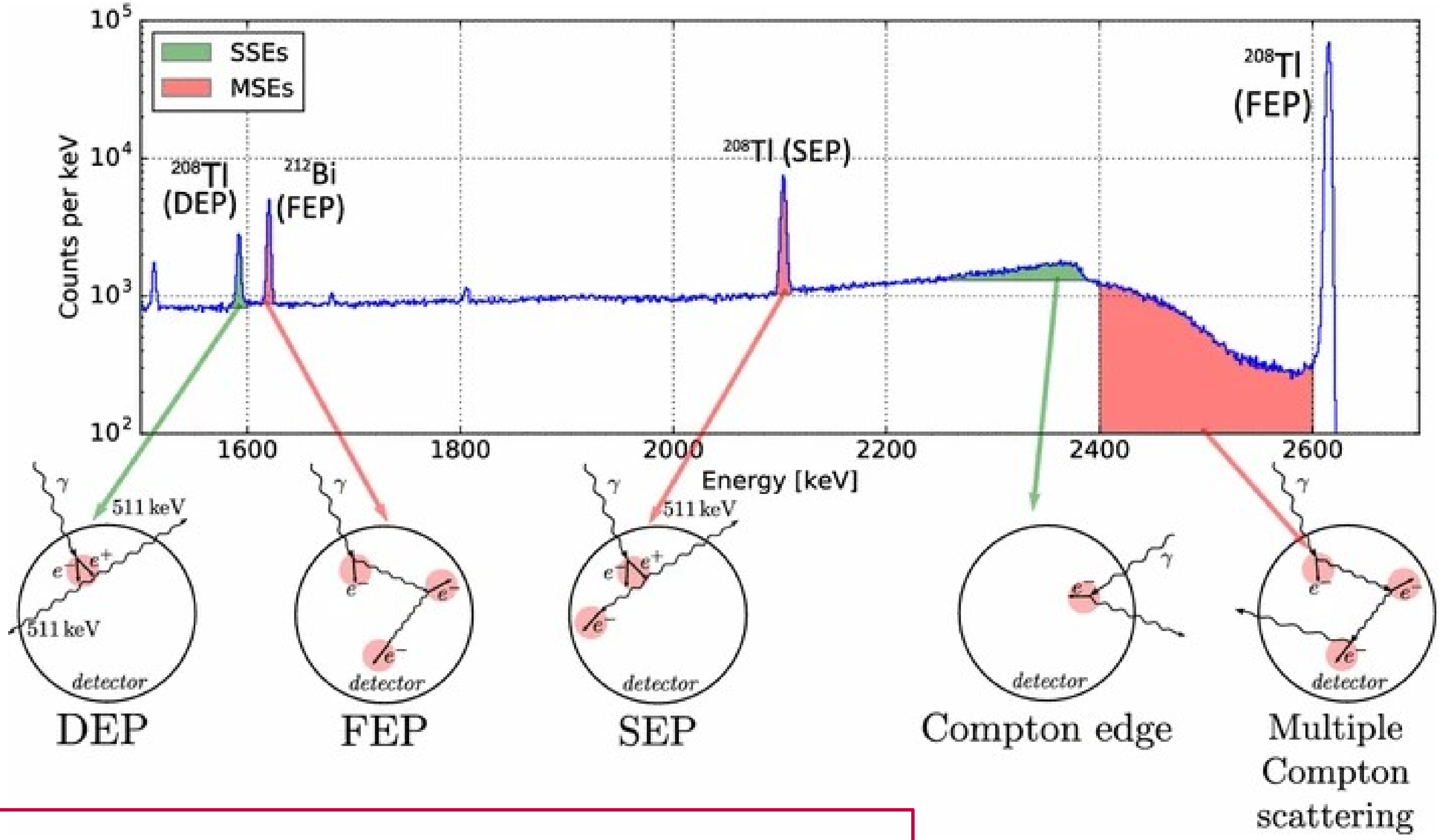
# Pulse Shape Analysis

SSE




MSE




**Keep in Mind:**

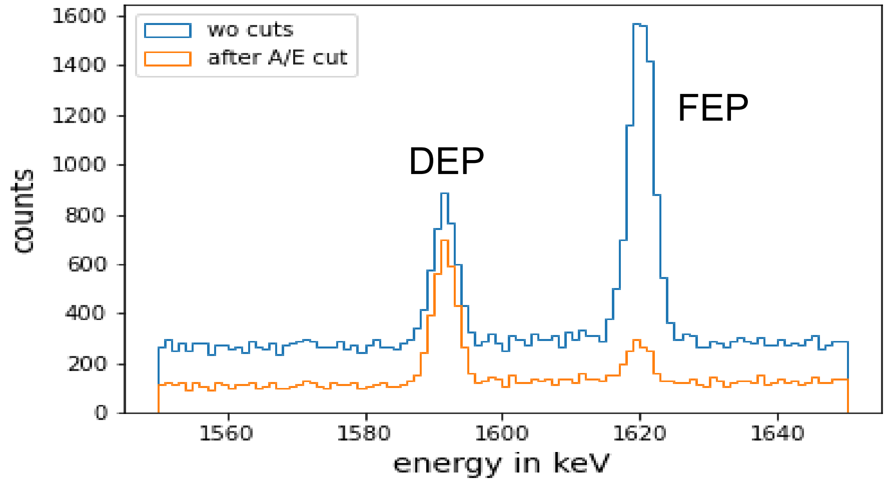
DEP: Signal like | FEP/SEP: bkg dominated | Continuum: Mixed

All peaks at different energies

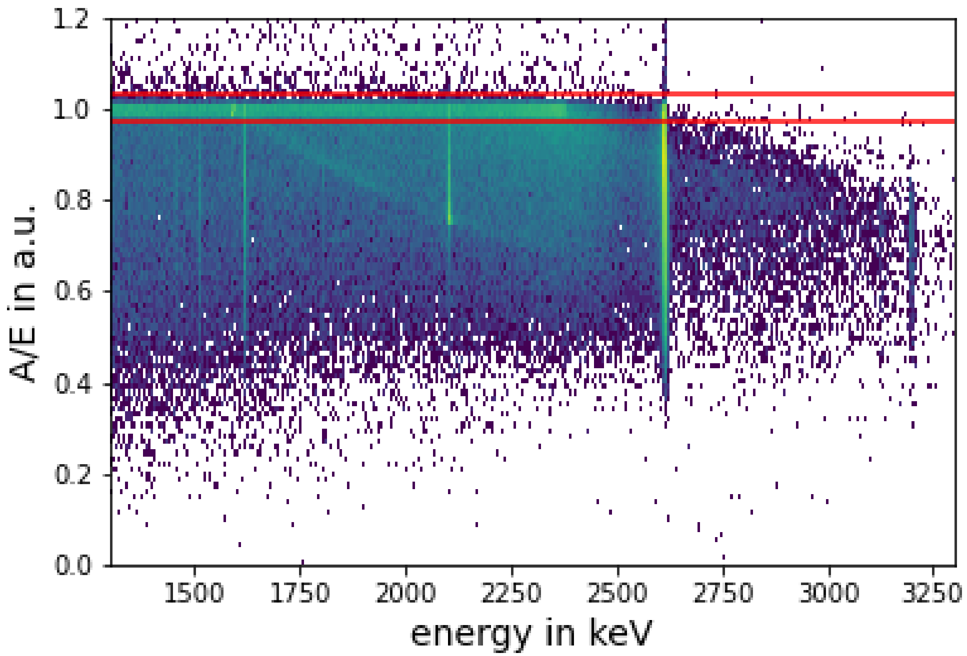
# Pulse Shape Analysis

**In General:**

1. Determine some kind of classifier
2. Set cut on classifier at 90% survival fraction in DEP

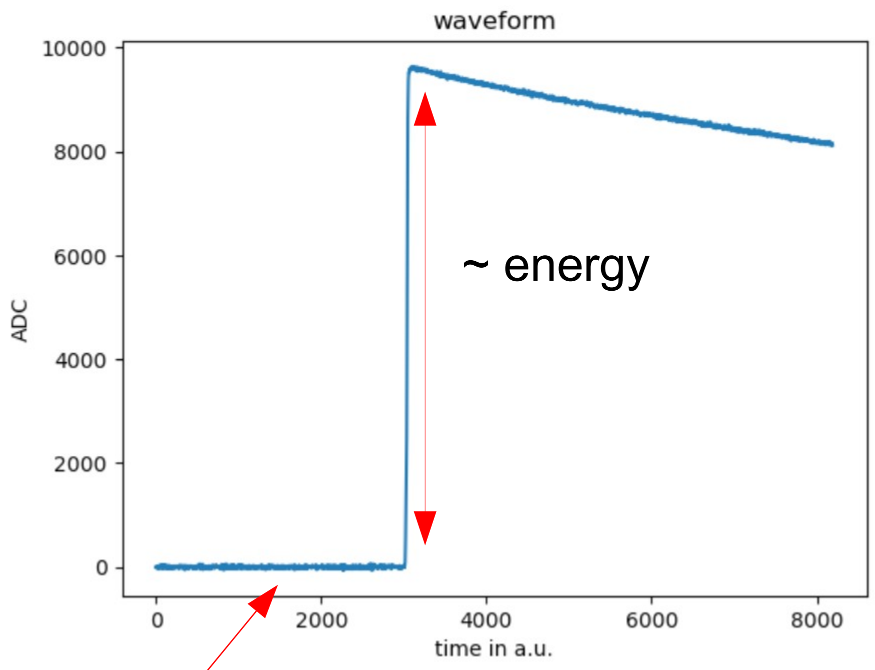



Classic A/E Analysis:

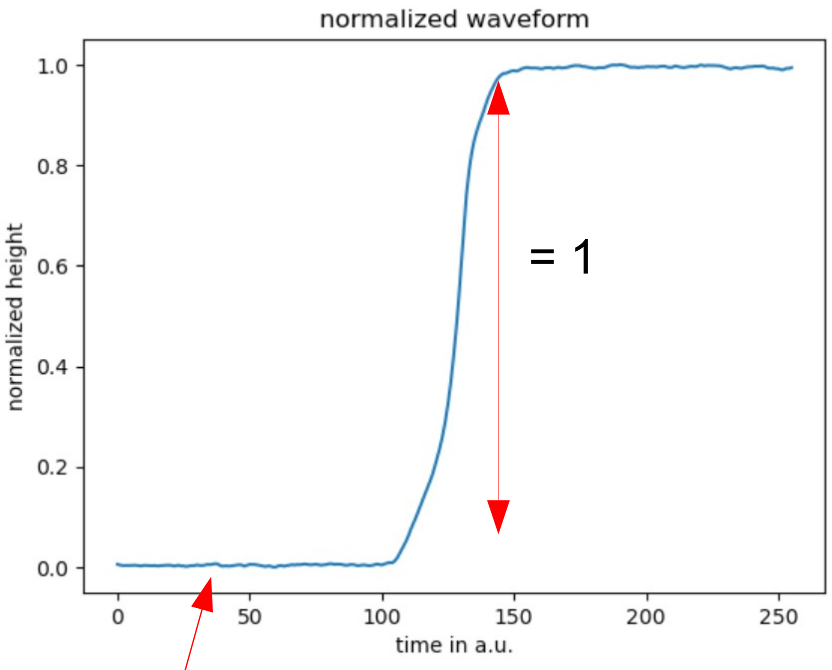



But does it work with Machine Learning too?

# Waveform Energy Dependence




Independent of energy




Energy dependent

**Keep in Mind:**
Baseline & tail contain energy information

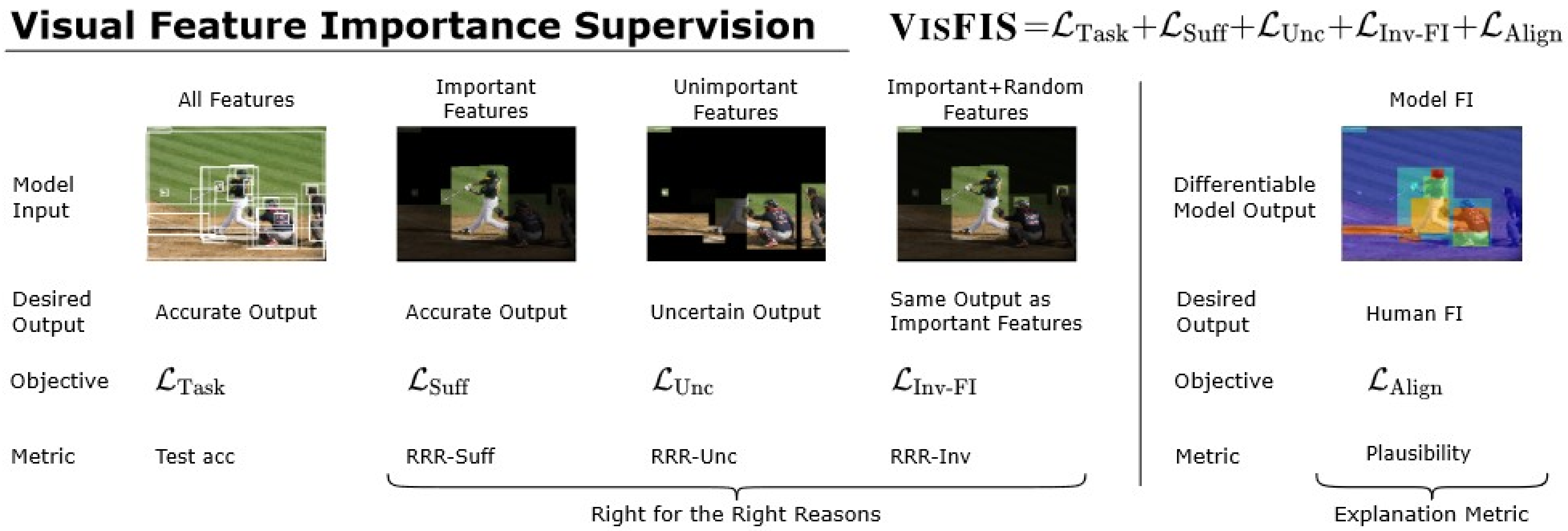
- Model can classify by using energy, not pulse shape
- Result shall be right for the right reasons

# Switch to the other Side: Machine Learning

What’s the idea behind Feature Importance Supervision (FIS?)

- Using human knowledge about important and unimportant features
- Lead the model to take right decisions

# Visual Feature Importance Supervision



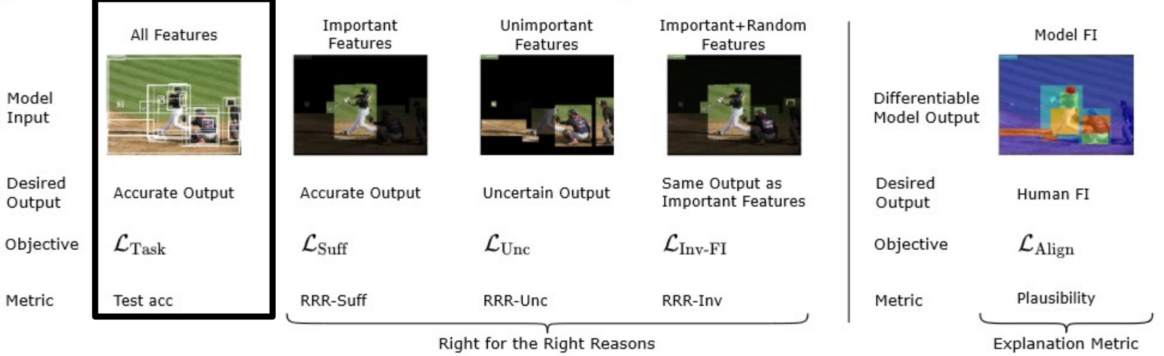

- VisFIS used for Visual Question Answering
- Use different variations of the input image
- Calculate a loss function for every variation
- Train with a combined loss function

https://arxiv.org/pdf/2206.11212.pdf

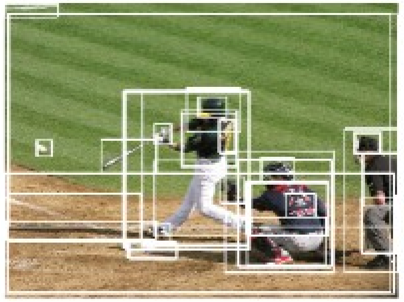

Zhuofan Ying, Peter Hase, and Mohit Bansal
Department of Computer Science
University of North Carolina at Chapel Hill

## Visual Feature Importance Supervision

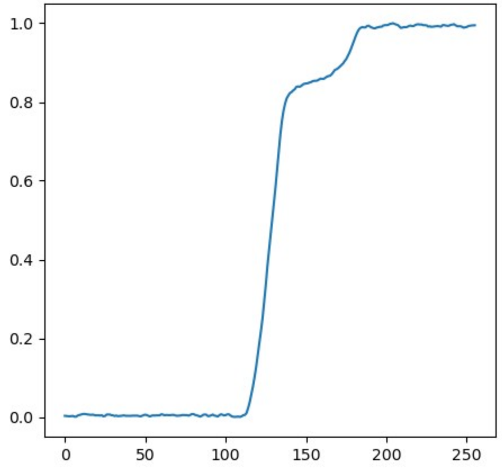
$\text{VisFIS} = \mathcal{L}_{\text{Task}} + \mathcal{L}_{\text{Suff}} + \mathcal{L}_{\text{Unc}} + \mathcal{L}_{\text{Inv-FI}} + \mathcal{L}_{\text{Align}}$




All Features



Adaption to Ge Signals

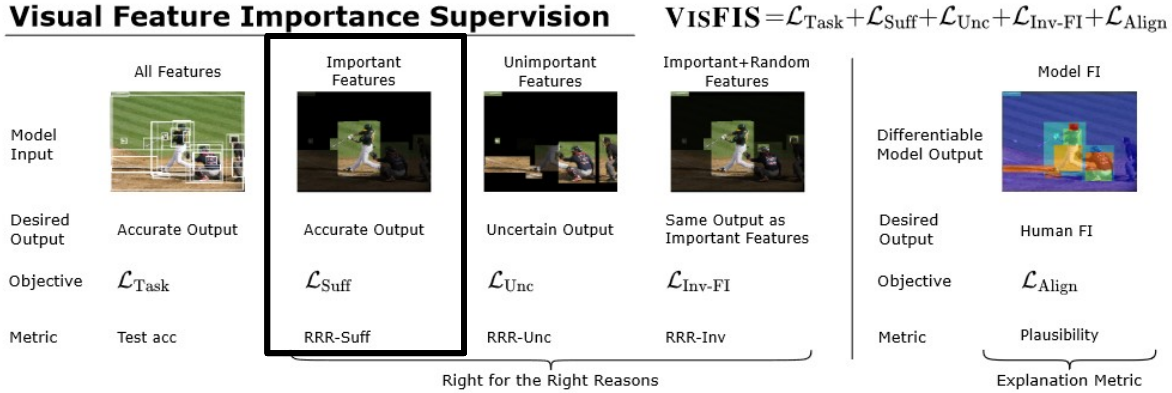



Accurate Output

$\mathcal{L}_{\text{Task}}$

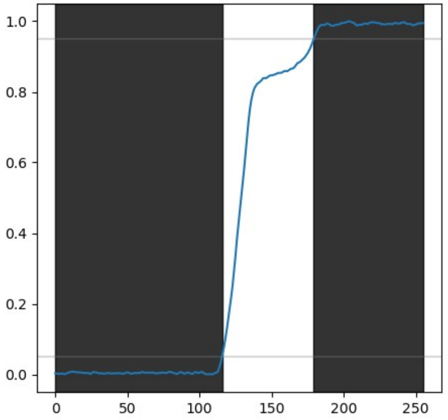
**Goal:** Given the full task input, the model shall return an accurate output.

**Method:** Feed into model and train model by using known label *y*.




Important Features

Adaption to Ge Signals

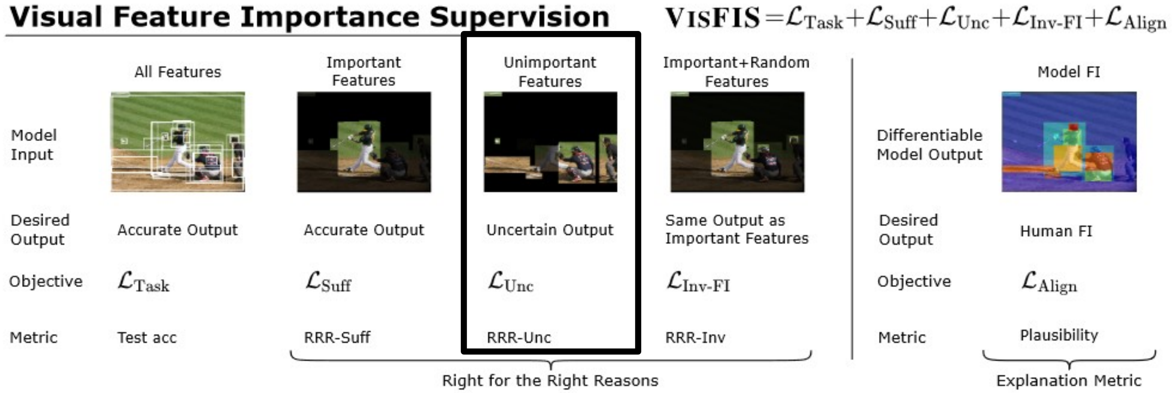



Accurate Output

$\mathcal{L}_{\text{Suff}}$

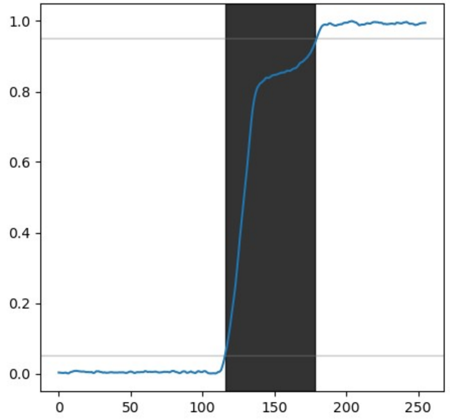
**Goal:** Subset of input containing the important features shall be sufficient to produce accurate output.

**Method:** Feed into model and train model by using known label *y* (as with the original input)




Unimportant Features

Adaption to Ge Signals




Uncertain Output

$\mathcal{L}_{\text{Unc}}$

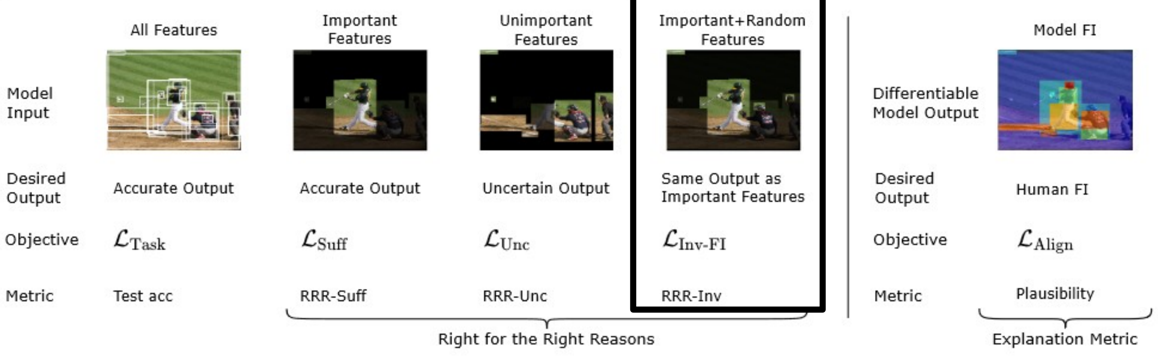
**Goal:** Subset of input containing just unimportant features shall result in total uncertainty.

**Method:** Feed into model and train model by using a random number as label.

$$\mathcal{L}_{\text{Unc}}(\theta, x, e) = \text{KL}\big(\text{Unif}(|\mathcal{Y}|), f_\theta(x_u)\big)$$

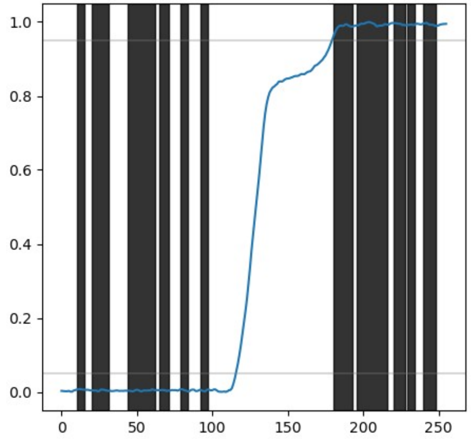
## Visual Feature Importance Supervision

$\text{VisFIS} = \mathcal{L}_{\text{Task}} + \mathcal{L}_{\text{Suff}} + \mathcal{L}_{\text{Unc}} + \mathcal{L}_{\text{Inv-FI}} + \mathcal{L}_{\text{Align}}$




Important+Random Features

Adaption to Ge Signals




Same Output as Important Features

$$\mathcal{L}_{\text{Inv-FI}}$$

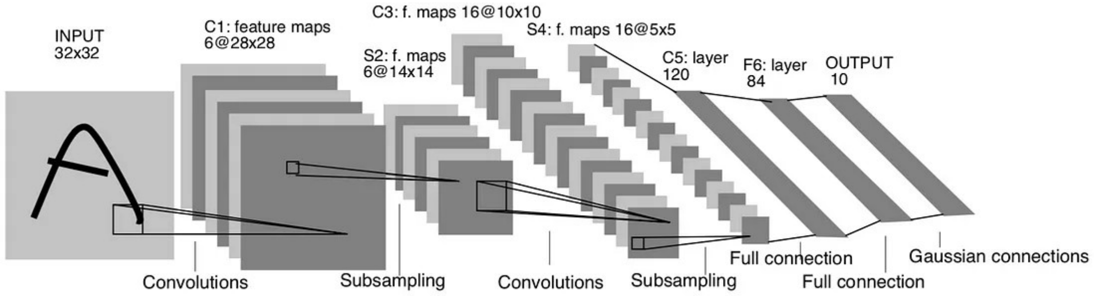
**Goal:** Model shall be invariant under important features + random sampling of unimportant features.

**Method:** Using input with important and unimportant features, train to same result as just with important input.

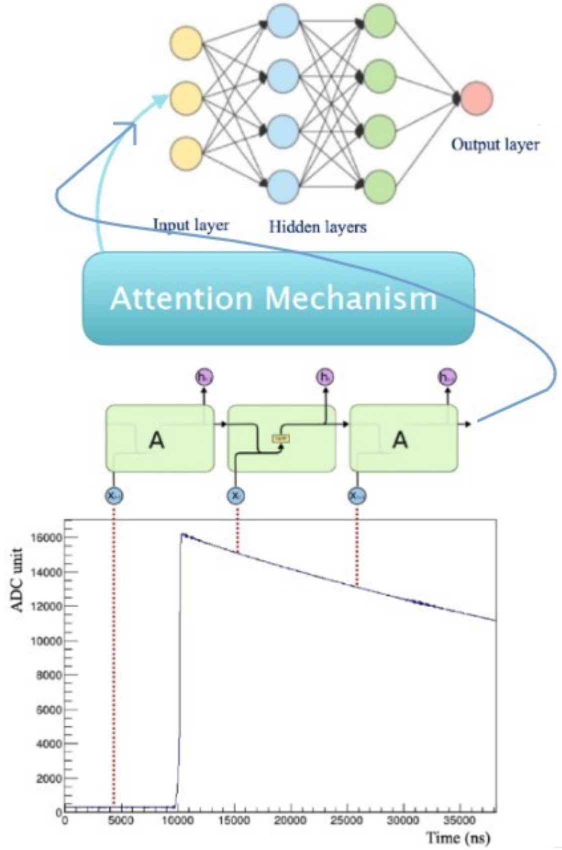
$$\mathcal{L}_{\text{Inv-DA}}(\theta, x, e, \mathcal{D}_u) =$$
$$\mathbb{E}_{u \sim \mathcal{D}_u} \text{KL}(f_\theta(x_e), f_\theta(x_{e \cup u}))$$

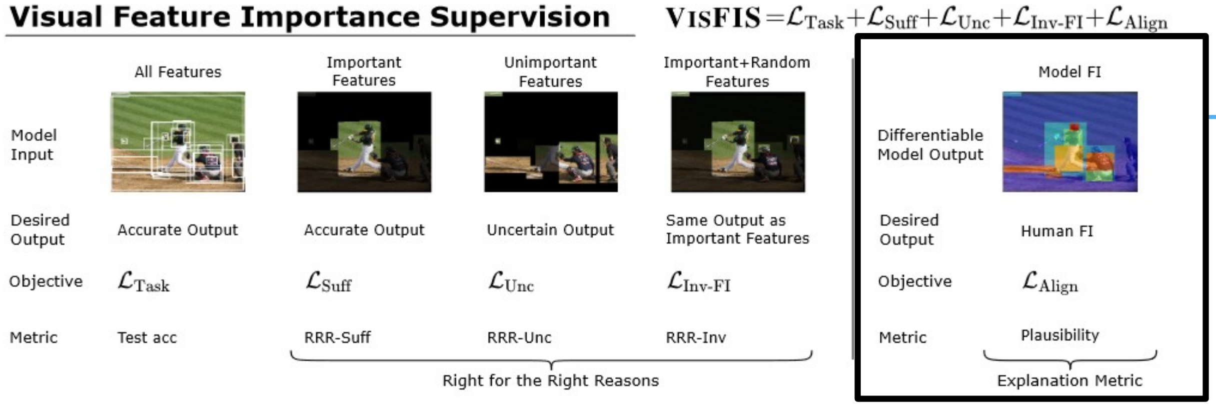
# CNN & RNN+attention model

- CNN:
  - Necessary to calculate vanilla gradient as explanation metric

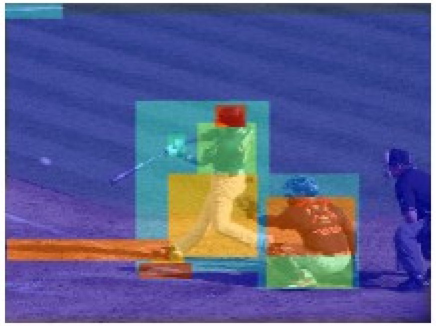



- RNN+attention:
  - Attention score intrinsic to model, leads direct to explanation metric

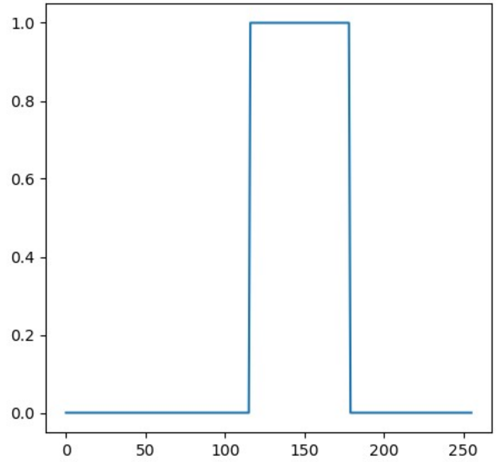





Model FI



Adaption to Ge Signals




Human FI

$$\mathcal{L}_{\text{Align}}$$

**Goal:** Want an alignment between human and model feature importance.

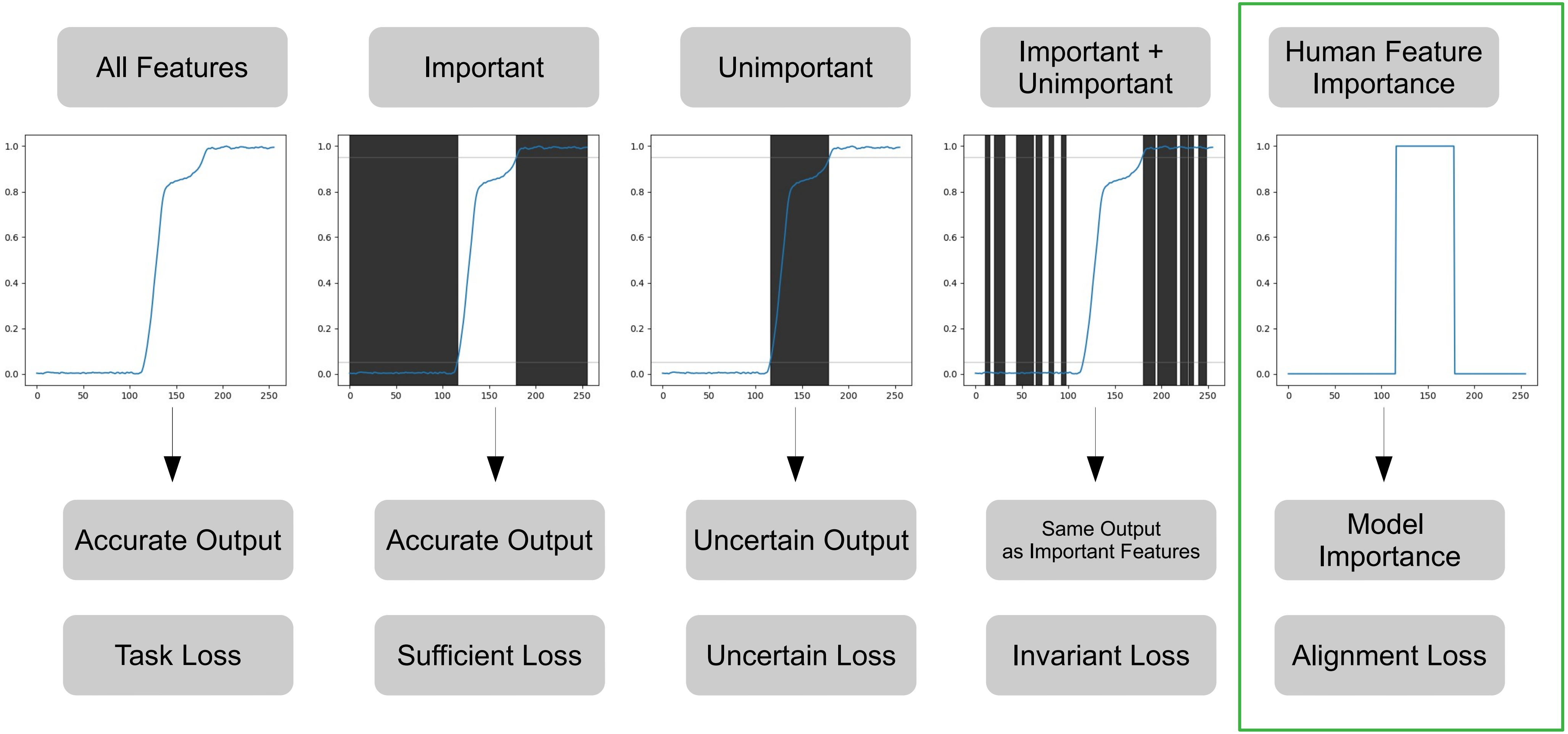
**Method:** Train model to result explanation metric *e* of the model having the same shape as human explanation ē.

**Explanation Metric:**
- Vanilla Gradient for CNN
- Attention score for RNN+attention

$$\mathcal{L}_{\text{align}}(\theta, x, e, \tilde{e}) = \text{cos-sim}(e, \tilde{e})$$

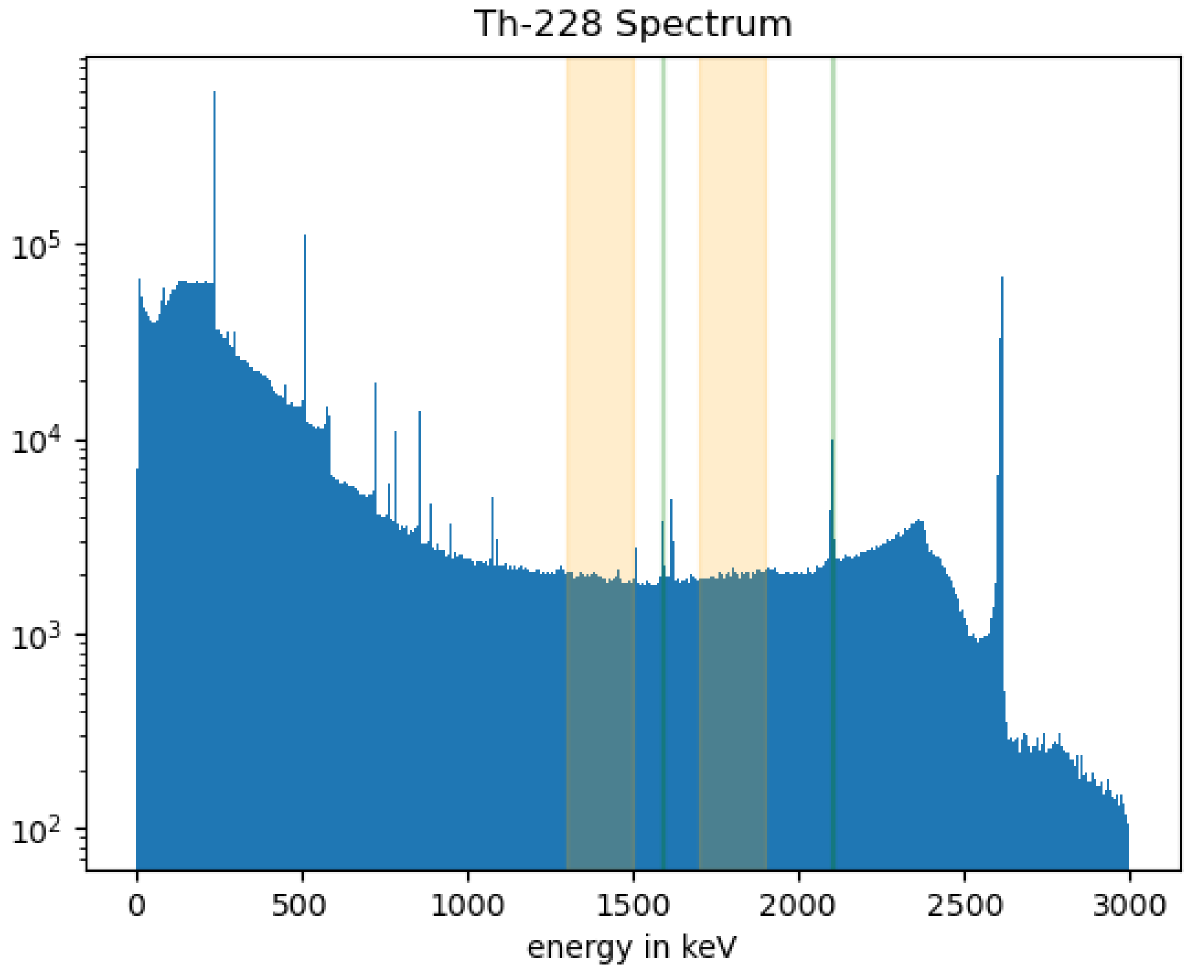
# Overview about FIS model

**About the explanation metric**




→ Add all together

# Training Setup

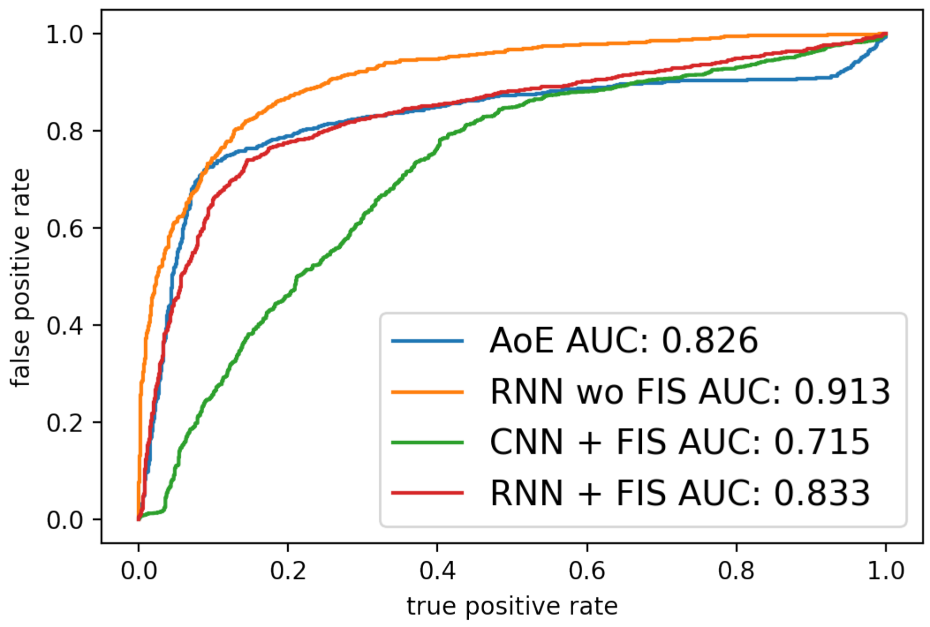



a) training peaks (DEP as signal type, SEP as bkg type)
b) check for energy dependence

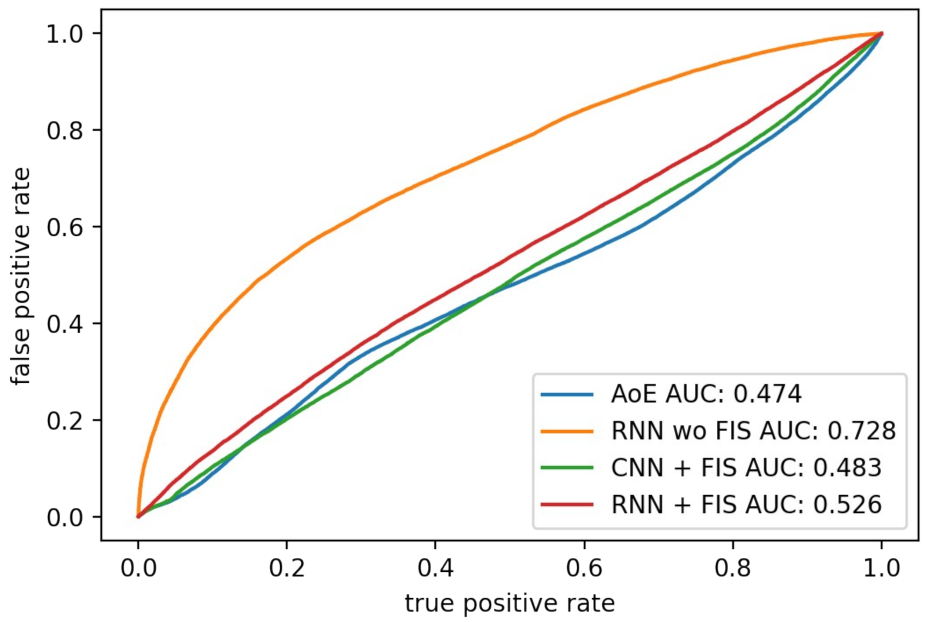
- Using characterization measurements
- Compare 4 Versions:
  - **A/E** as non-ML analysis
  - **RNN wo FIS**
  - **CNN** as a basic model
  - **RNN+attention** as an advanced model

# Classification power

### Model performance




### Check for energy dependence

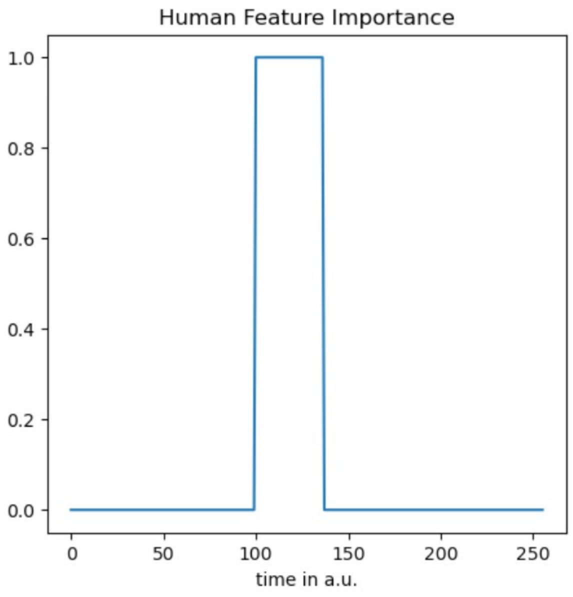


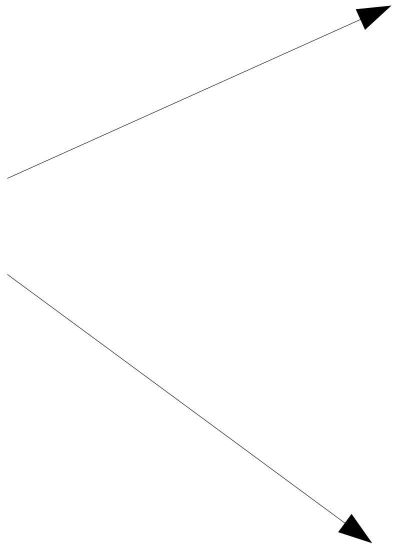

→ models with FIS are energy independent

→ RNN + FIS seem to perform very well

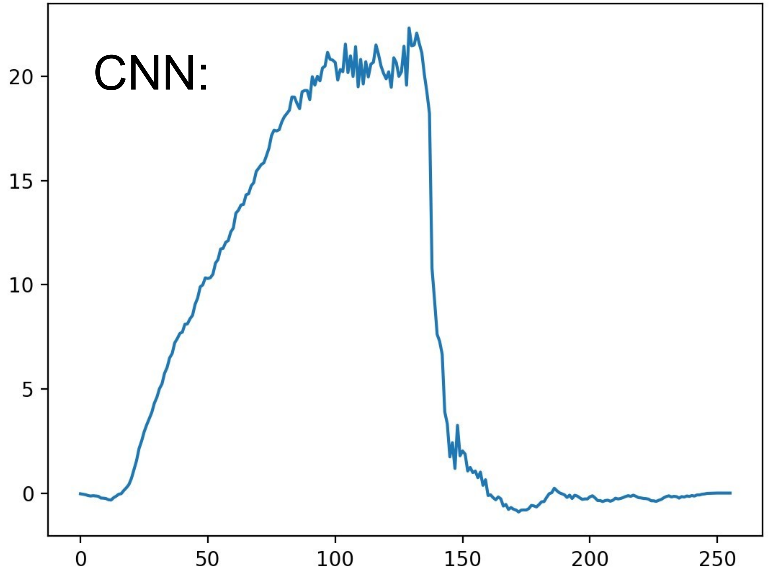
# Training results of FIS

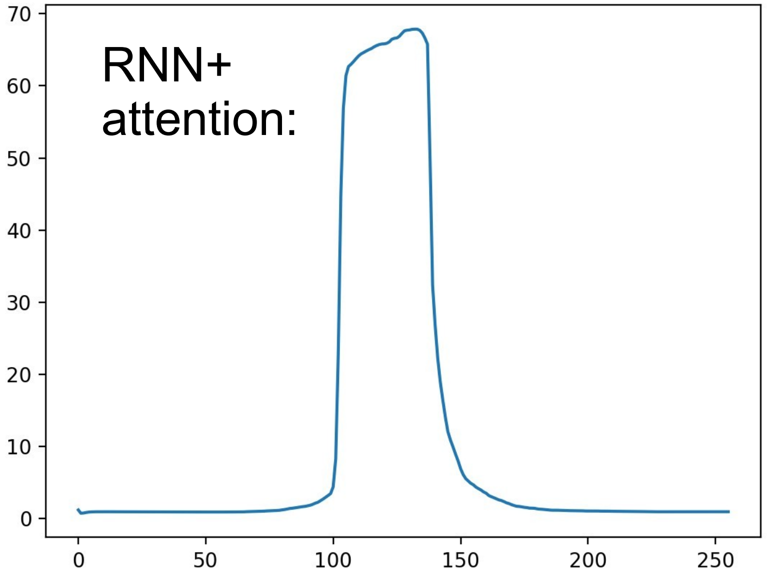
Human Input:



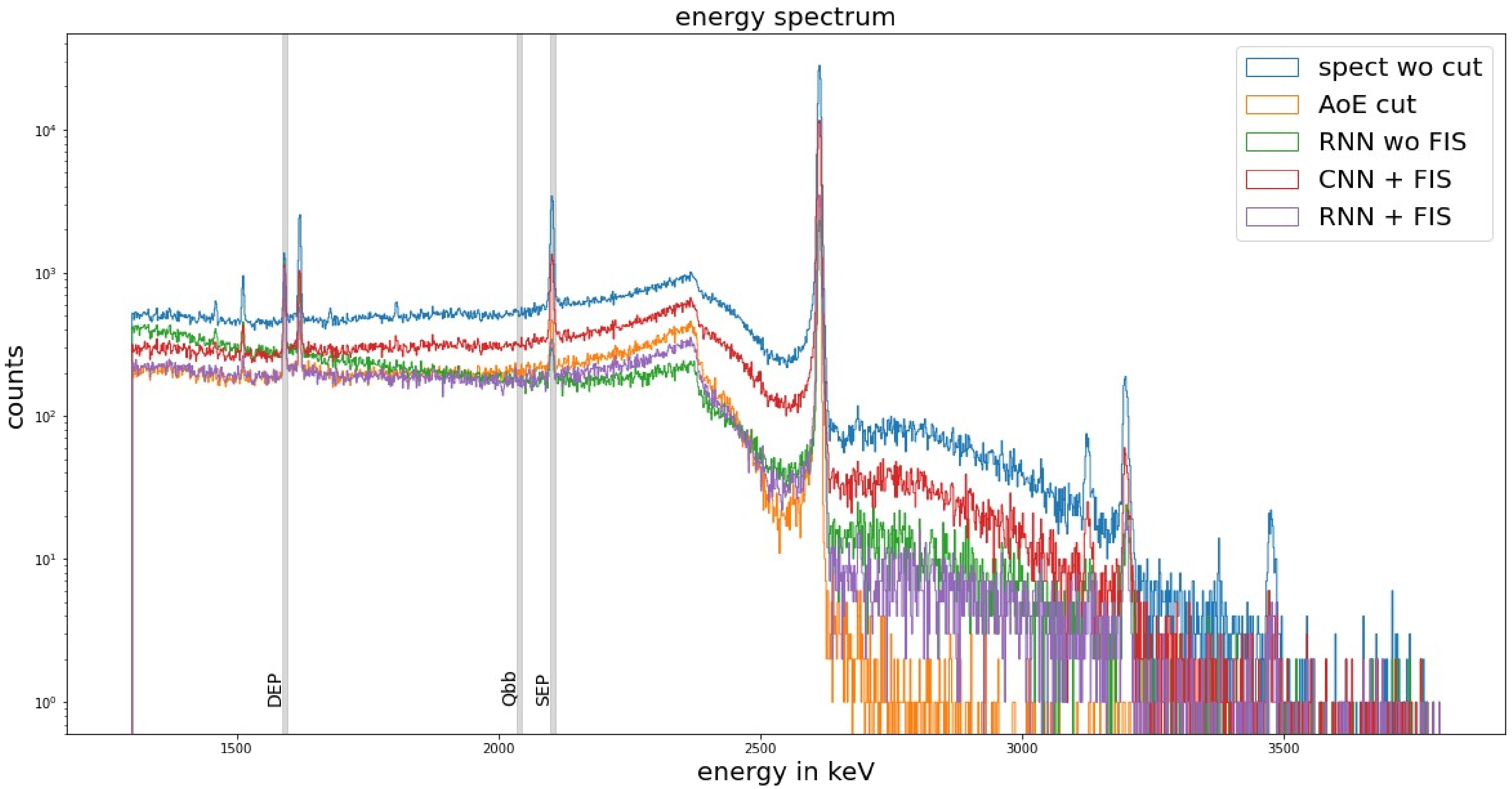



Machine Output:



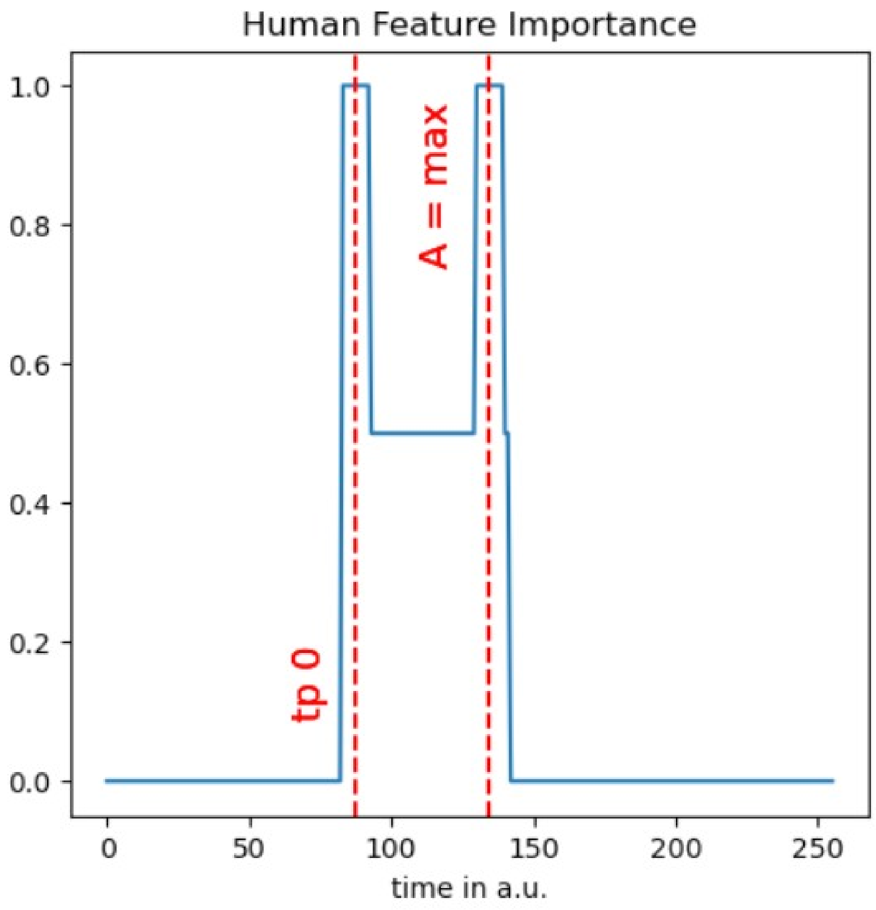
# Spectrum after cuts



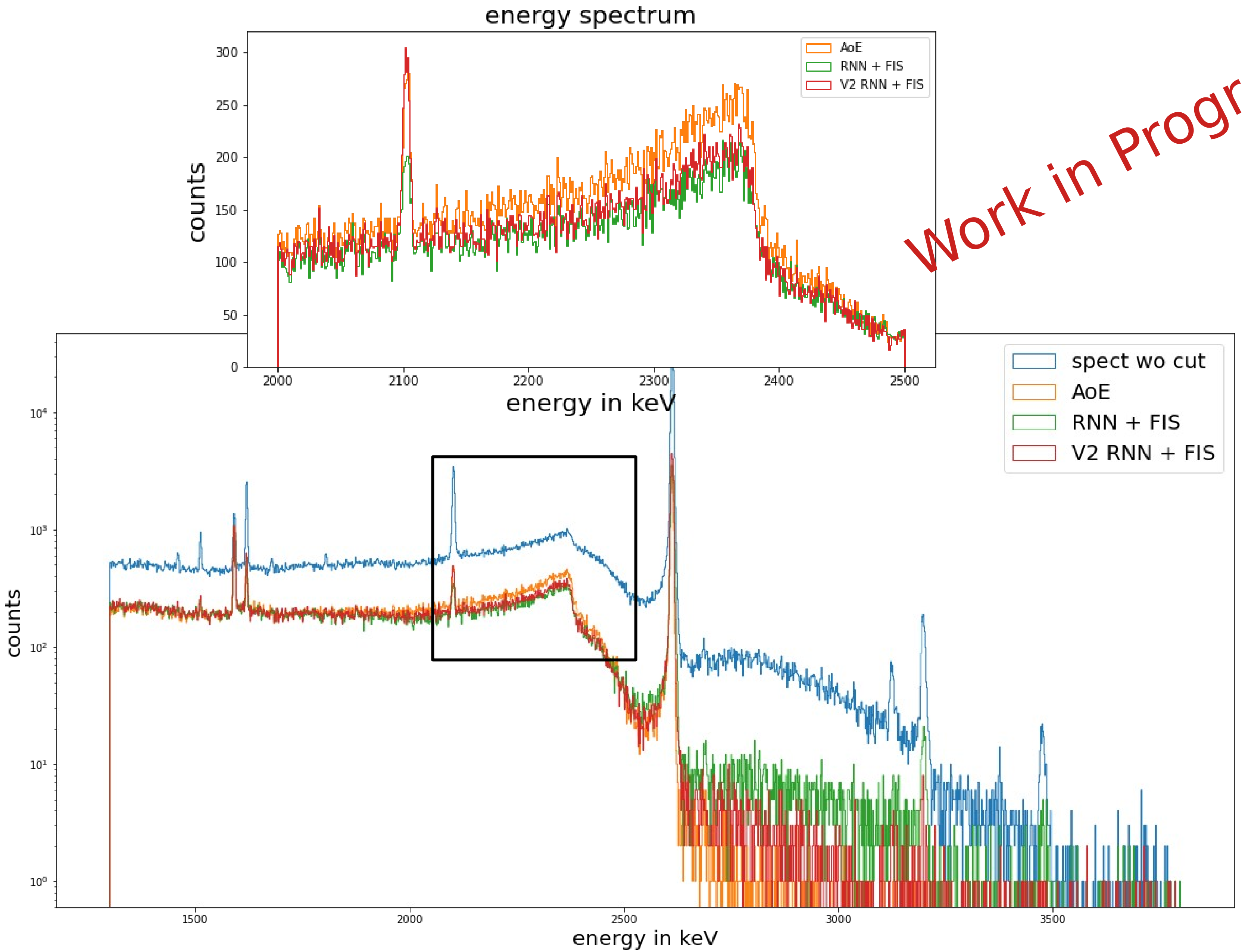

| peak | DEP | SEP | Qbb |
|---|---|---|---|
| type | signal | bkg | mixed |
| A/E | 90% | 7% | 29% |
| RNN | 90% | 5% | 33% |
| CNN +FIS | 90% | 36% | 60% |
| RNN +FIS | 90% | 6% | 33% |

# Further possibilities by FIS

- Adding more physical knowledge
  - Starting point
  - Point of maximal A
  - ...?
- 2nd version of Human FI
- Other versions possibles



# Further possibilities by FIS

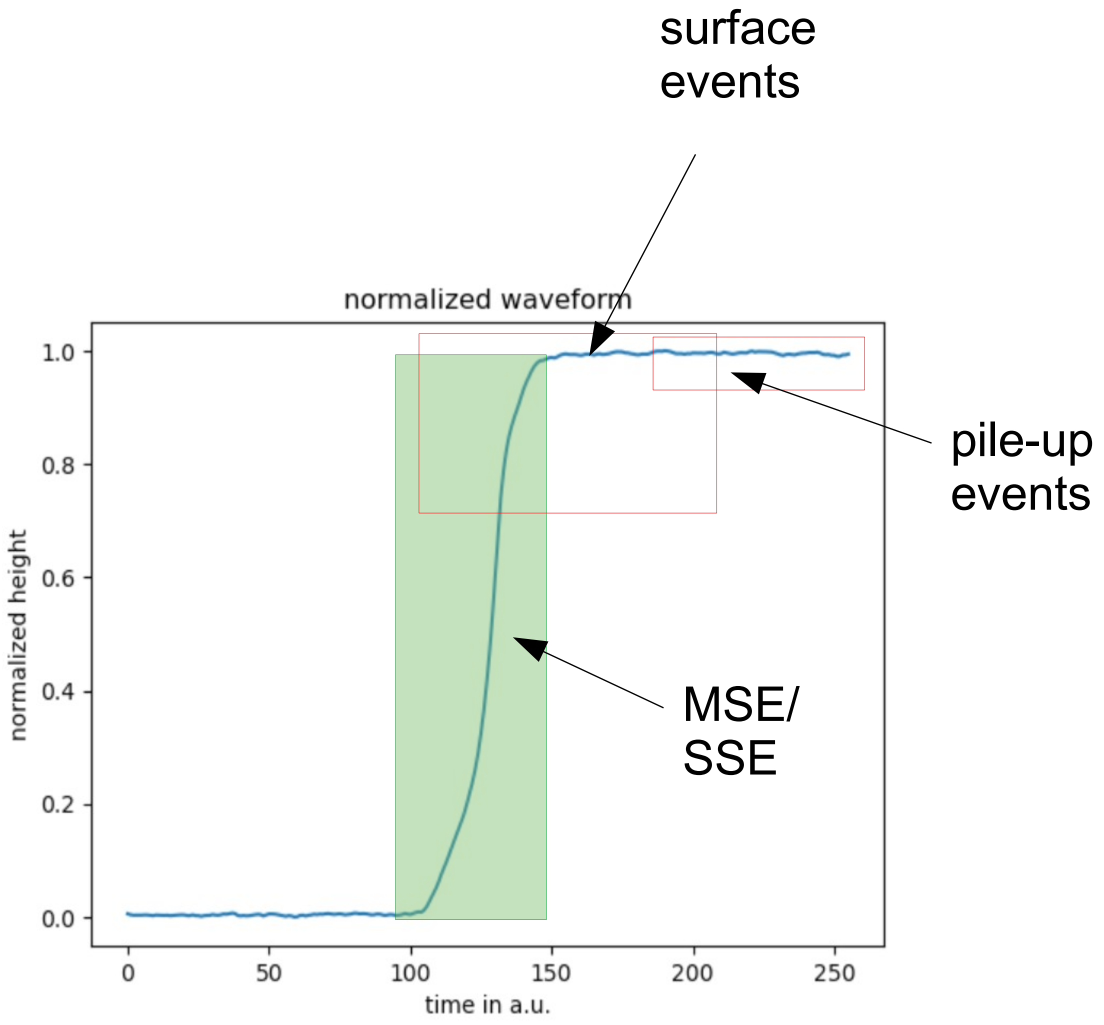



Work in Progress!

| peak | DEP | SEP | Qbb |
|---|---|---|---|
| type | signal | bkg | mixed |
| A/E | 90% | 7% | 29% |
| RNN +FIS | 90% | 6% | 33% |
| V2 RNN +FIS | 90% | 10% | 35% |

# Further possibilities by FIS

- Use FIS for other PSA tasks
- Focussing on different part of the waveform, depending on the task
  - Surface events (alphas, betas)
  - Pile-up events
  - Position reconstruction
  - ...



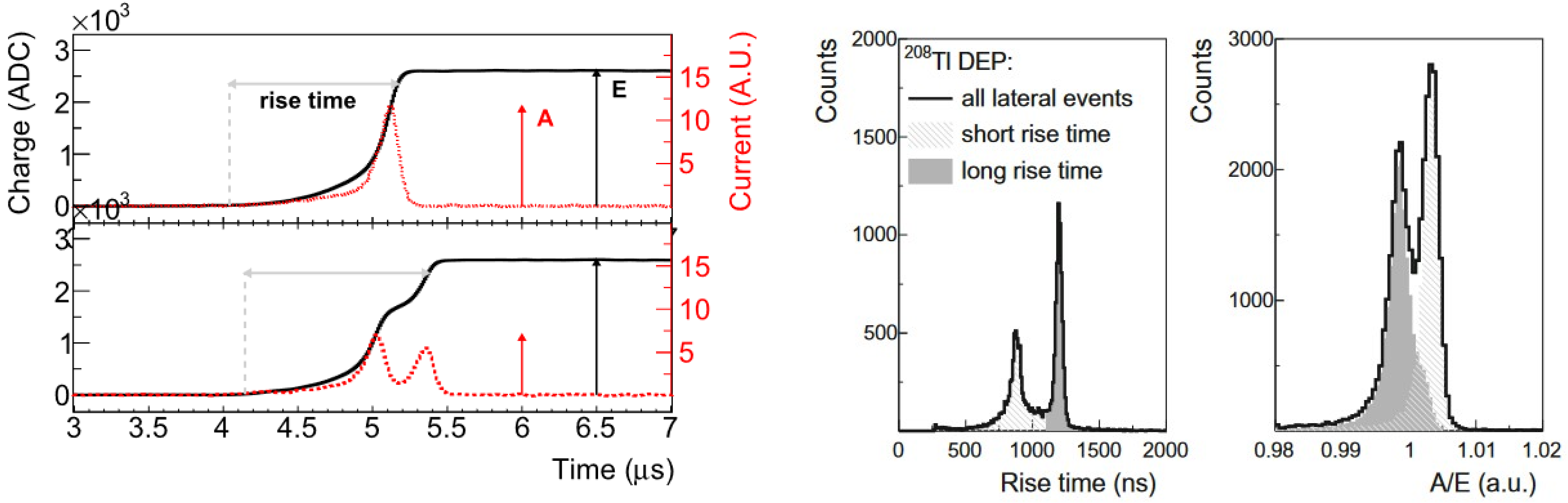
# Summary

- Possible to remove energy dependence
- RNN better than CNN
- RNN+FIS close to A/E
- Still some differences to investigate
- Further steps:
  - Check behaviour in compton shoulder
  - Investigate model performance on low-background physics data
  - FIS for other classes of events

More Questions? Klick here to send an eMail

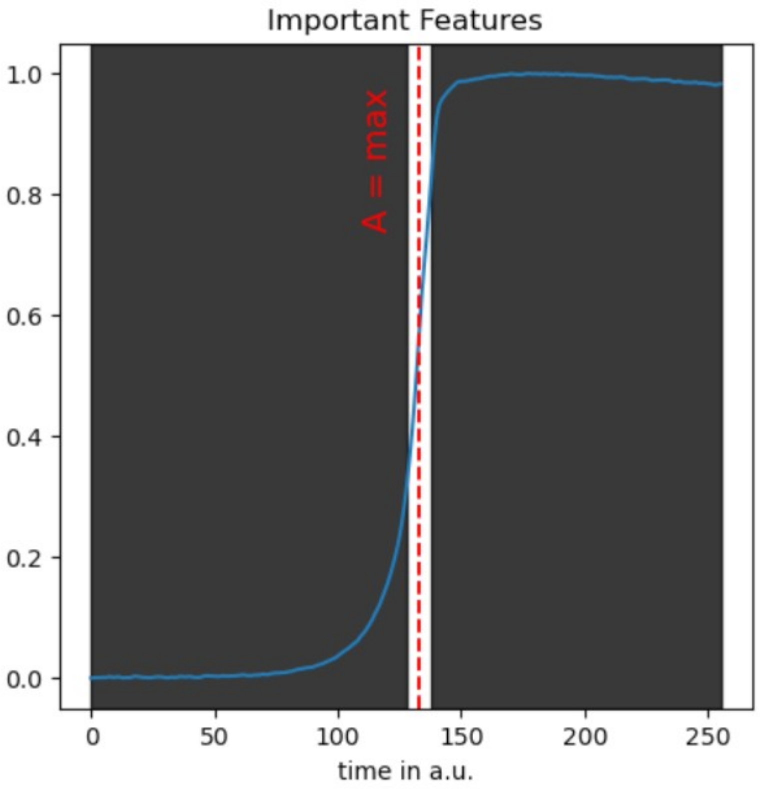
# Backup

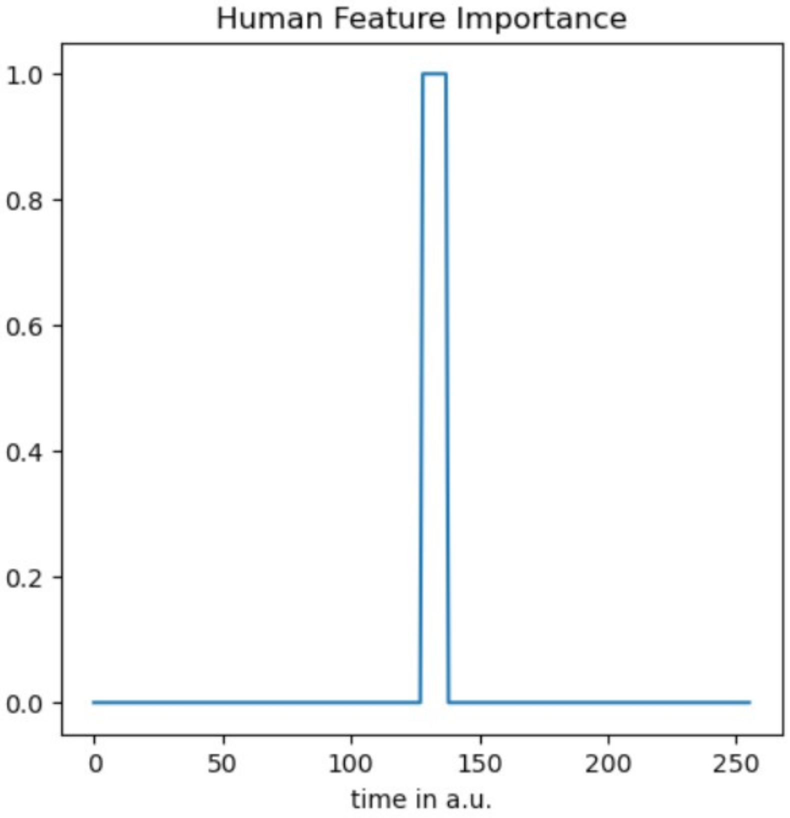
# Rise time



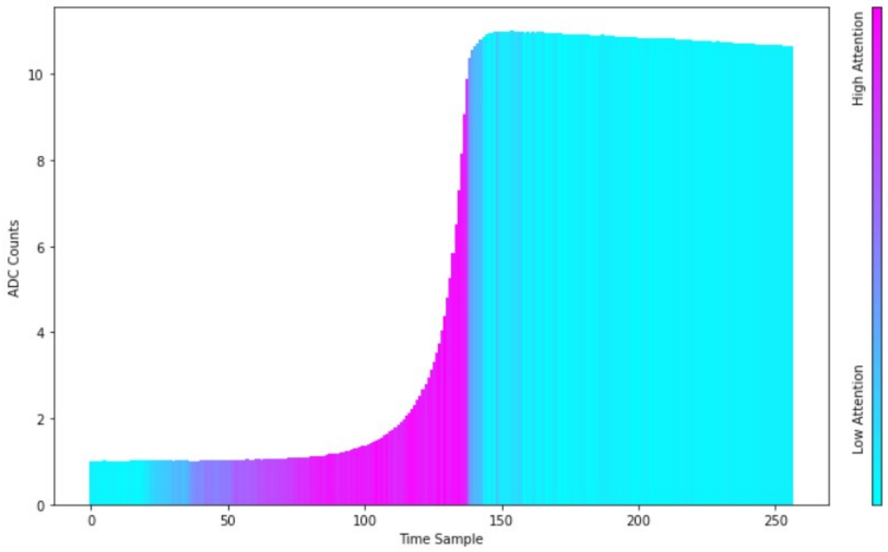

- https://doi.org/10.1140/epjc/s10052-021-09184-8
- Depends on position inside detector (short RT close to p+ contact)

# Backup

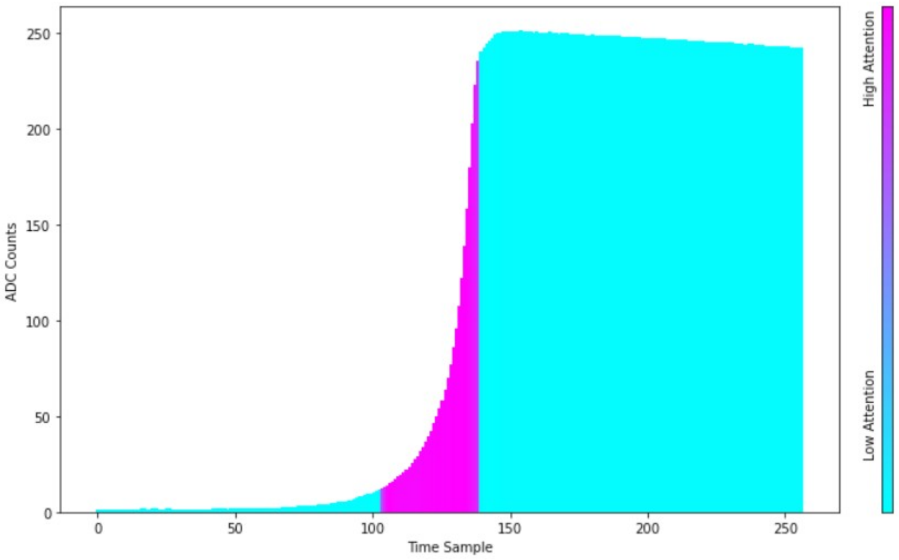

Important Features
A = max
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
time in a.u.


Human Feature Importance
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
time in a.u.

# Backup



High Attention
Low Attention
ADC Counts
Time Sample

High Attention
Low Attention
ADC Counts
Time Sample