CEPC AHCAL PID

Use Artificial Neural Network

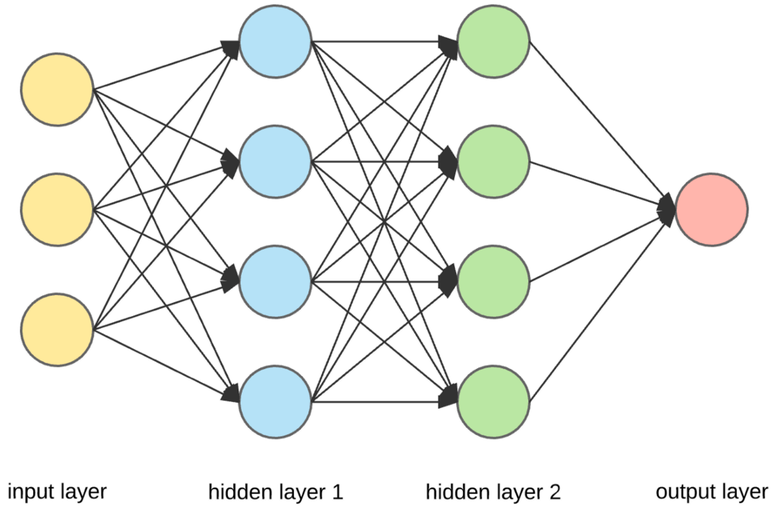
Siyuan SONG, Zhen WANG, Zixun XU

Content

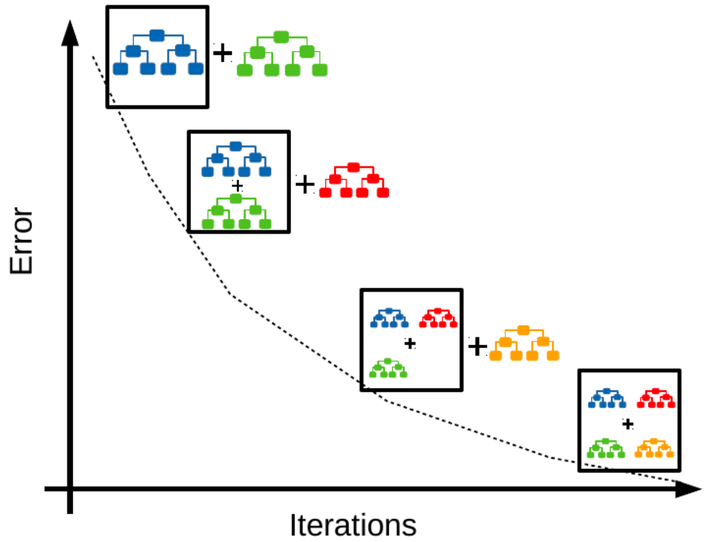
1. Brief ANN principle.
2. Datasets introduction.
3. Preliminary results.
4. ANN vs BDT comparison.

Technical route

- **A**rtificial **N**eural **N**etwork:
  Siyuan SONG



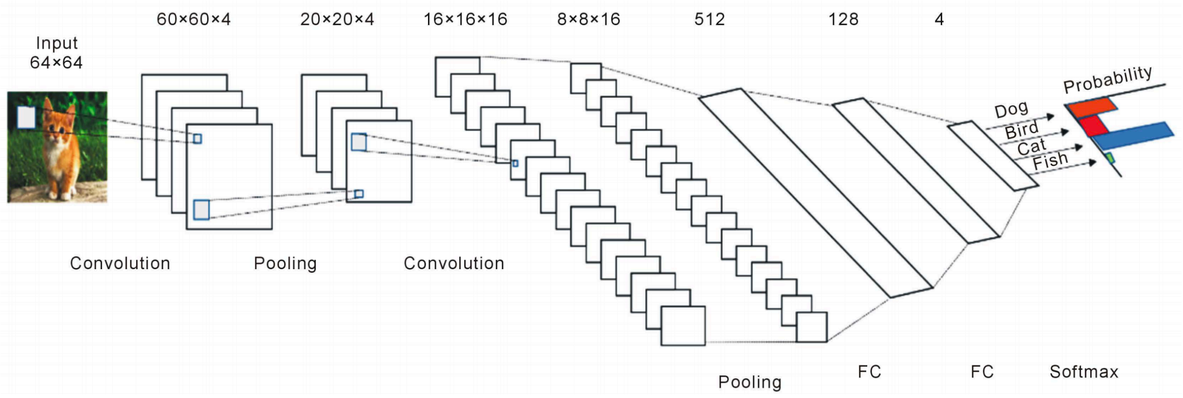
- **B**oosted **D**ecision **T**ree:
  Zhen WANG



ANN Fundamental principle

- Need an Image (Tensor) and a label (Tensor).
- Output could be a normalized Tensor storing several probabilities.
- Net actually is a very complicated function:
  - Output = Net(image, net parameters)
- Lose function would guide optimizing parameters of the net in Gradient Descent way.
  - Loss = L(output, label)

Net



Forward: get classification probability.

Backward: optimize net parameters.

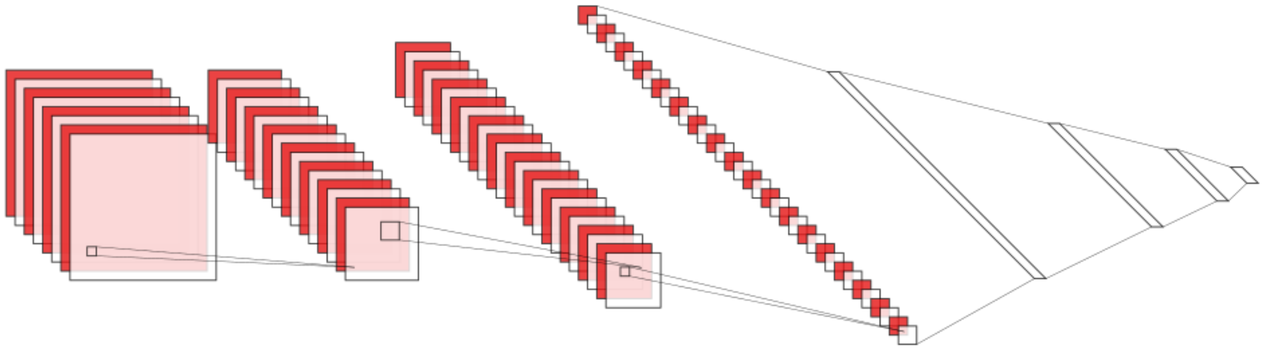
$$L_{\mathrm{CE}} = -\sum_{i=1}^{n} t_i \log(p_i), \quad \text{for n classes,}$$

where t_i is the truth label and p_i is the Softmax probability for the i^{th} class.

Loss function

ANN

- Input: tensor(40, 18, 18), energy deposits.
- Output: tensor(3,), probability.
- 6,660,723 trainable paramaters in Net.



Input Tensor

Output Tensor

Feature Layer

Classification Layer

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1          [-1, 80, 16, 16]          28,880
       BatchNorm2d-2          [-1, 80, 16, 16]             160
            Conv2d-3         [-1, 160, 14, 14]         115,360
       BatchNorm2d-4         [-1, 160, 14, 14]             320
            Conv2d-5          [-1, 320, 5, 5]         461,120
       BatchNorm2d-6          [-1, 320, 5, 5]             640
            Conv2d-7          [-1, 480, 3, 3]       1,382,880
       BatchNorm2d-8          [-1, 480, 3, 3]             960
            Linear-9                [-1, 960]       4,148,160
      BatchNorm1d-10                [-1, 960]           1,920
           Linear-11                [-1, 480]         461,280
      BatchNorm1d-12                [-1, 480]             960
           Linear-13                [-1, 120]          57,720
           Linear-14                  [-1, 3]             363
================================================================
Total params: 6,660,723
Trainable params: 6,660,723
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.05
Forward/backward pass size (MB): 1.00
Params size (MB): 25.41
Estimated Total Size (MB): 26.46
----------------------------------------------------------------
```

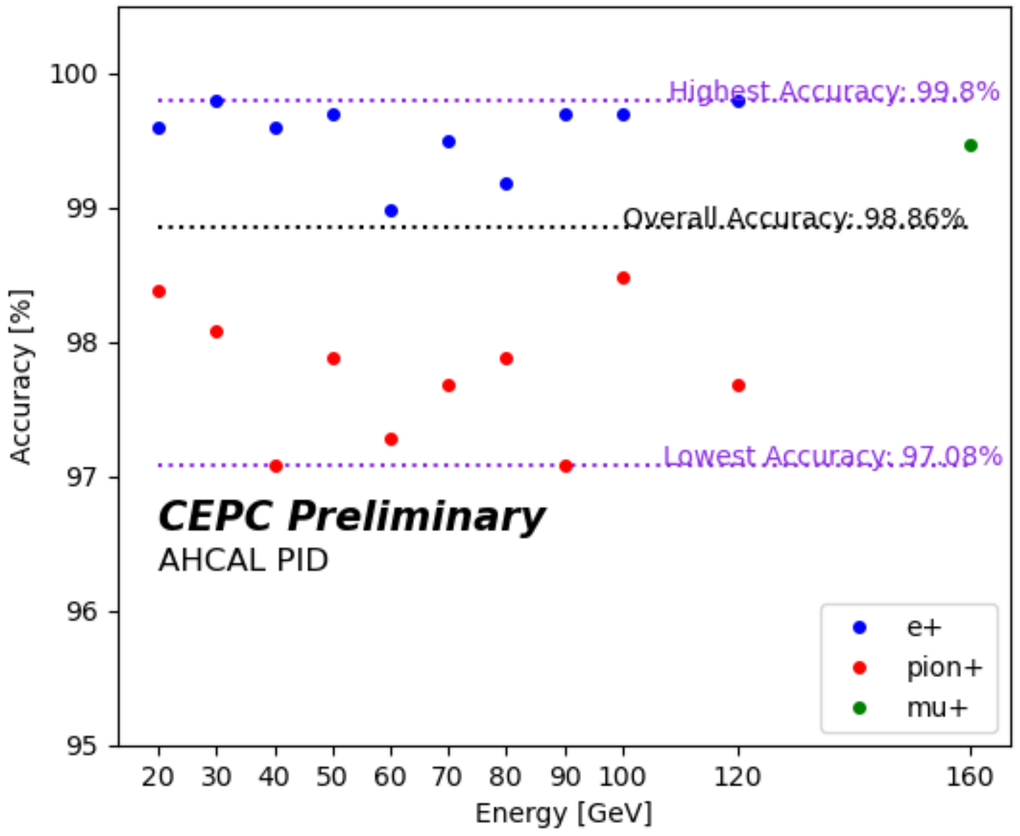
Data preparation

Particle\E (GeV)	20	30	40	50	60	70	80	90	100	120	160	Total	Finished
mu+											100k	100k	100k
e+	10k	10k	10k	10k	10k	10k	10k	10k	10k	10k		100k	100k
pion+	10k	10k	10k	10k	10k	10k	10k	10k	10k	10k		100k	100k

- Train set: Validation set: Test set = 8 : 1 : 1.
- Source: Geant4 simulation data, provided by Zixun XU.

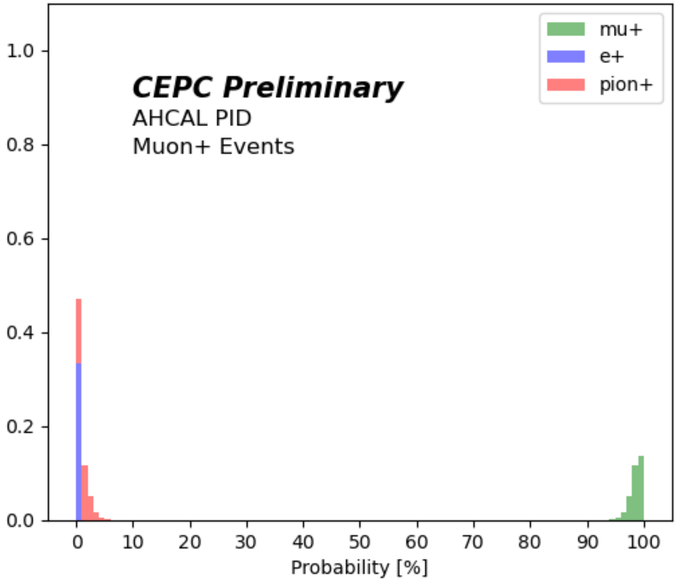
Preliminary results

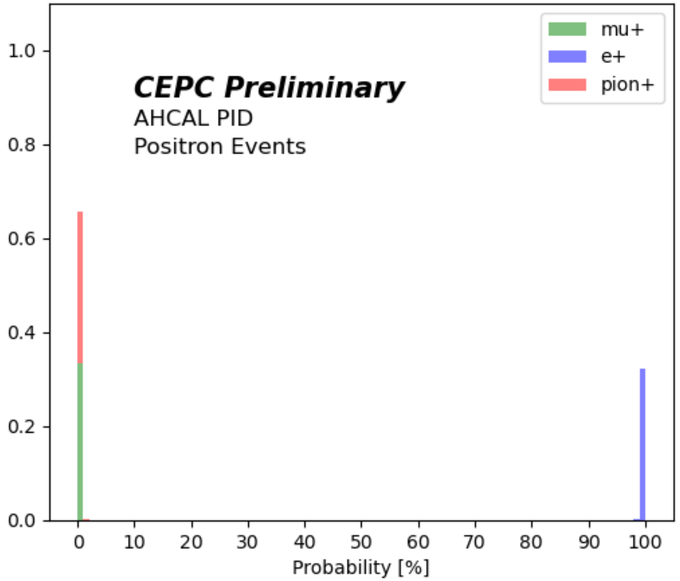
- Reached an encouraging overall accuracy: 98.86%.
- Each energy point: 1000 Events.
- e+ and muon+ can be classified with higher accuracy.
- pion+ events identification get a relative lower accuracy, but still > 97.08%.

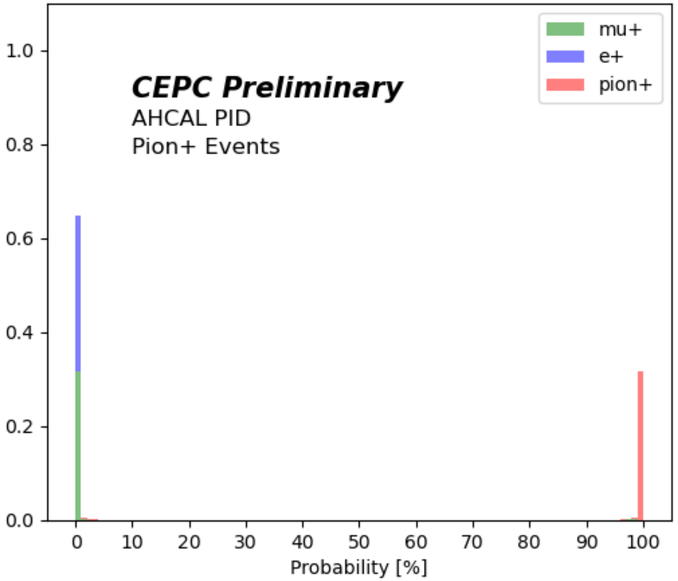


Preliminary results

- Net output probability distribution.
- Each kind of particle: 10,000 Events.
- From left to right, incident particles are Muon+, Positron+, and Pion+.
- The distinction is obvious.

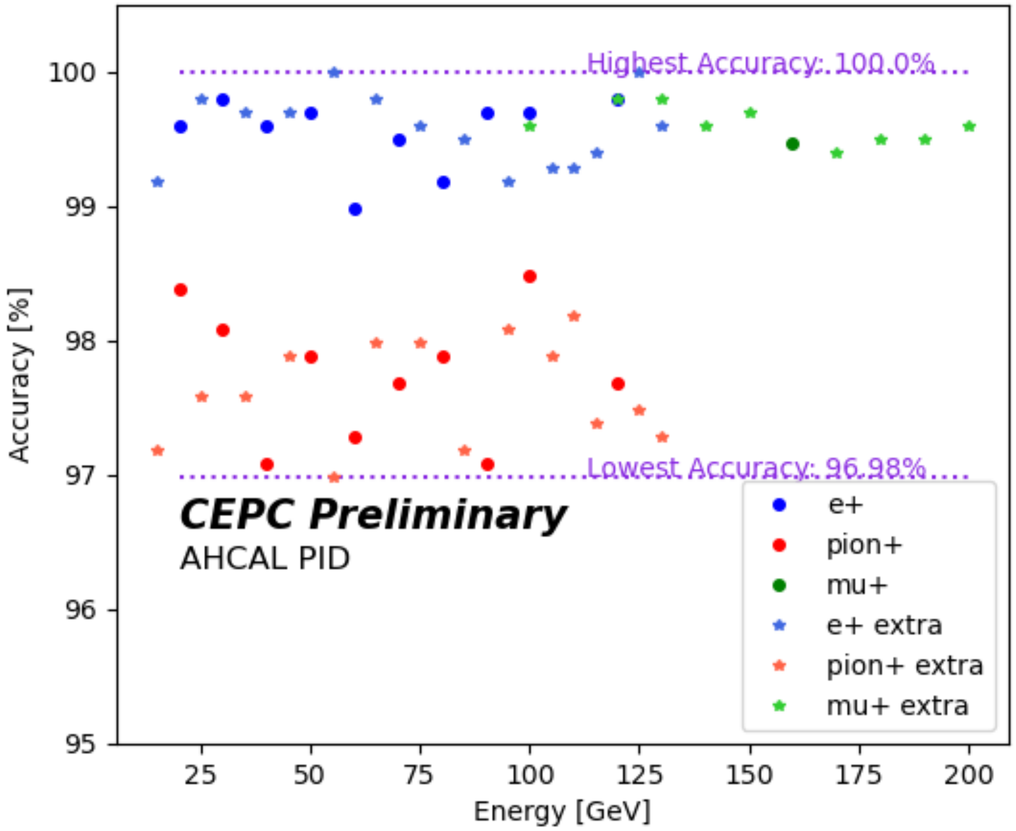






Preliminary results

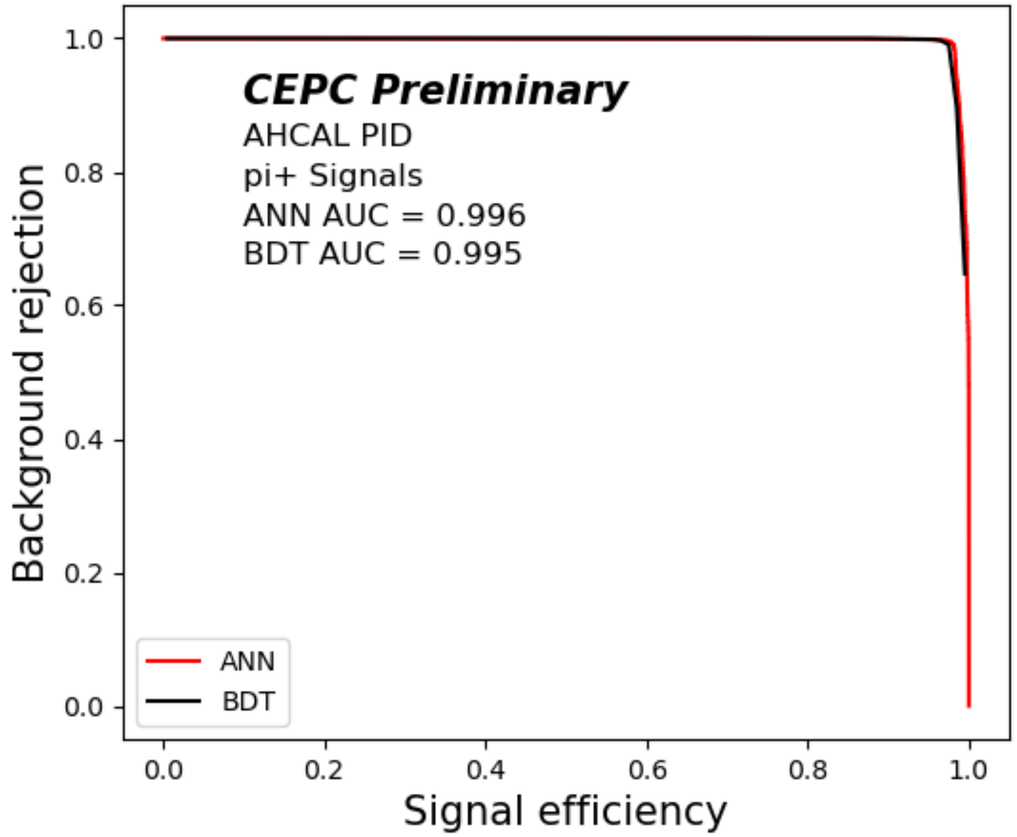
- Include muon+, e+, pion+ events with extra energy.
- Each energy point: 1000 Events.
- This net could still identify “unseen” particles.
  - e+: 15, 25, 35, 45, 55, 65, 75, 85, 95, 105, 110, 115, 125, 130 GeV.
  - pion+: 15, 25, 35, 45, 55, 65, 75, 85, 95, 105, 110, 115, 125, 130 GeV.
  - muon+: 100, 120, 130, 140, 150, 170, 180, 190, 200 GeV.



Preliminary results

ANN vs BDT

- Their ROC curve is very close in Pi+ signals.
- AUC stands for "Area under the ROC Curve.” Classifier is better if AUC is closer to 1.
- ANN AUC = 0.996 , when signal: Pi+, background: e+ & mu+ .



BDT ROC is provided by Zhen WANG.

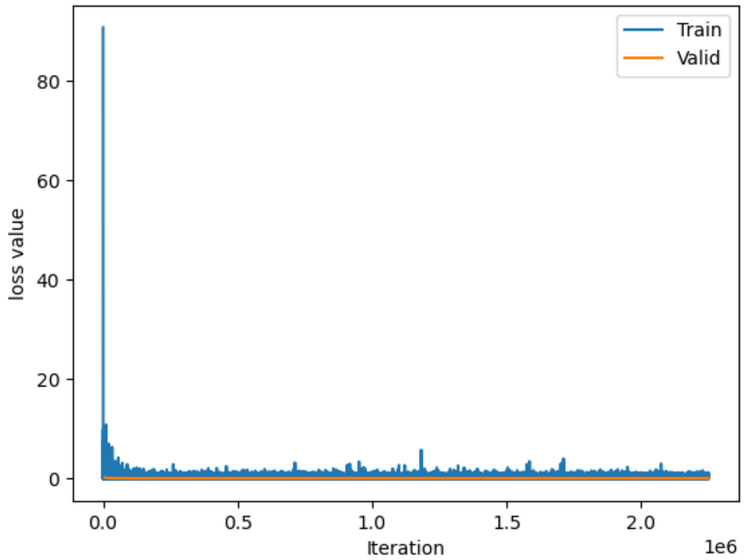
Ref: Park S H, Goo J M, Jo C H. Receiver operating characteristic (ROC) curve: practical review for radiologists[J]. Korean journal of radiology, 2004, 5(1): 11-18.

Future plan & Conclusion & Reflection

- Overall accuracy reached 98. 86%, ANN $\mathrm{AUC}_{\pi_{\mu e}}$ reached 0.996.
- Tune hyper-parameters; Optimize ANN structure.
- Compare with BDT method.
- Add ANN PID module into EventDisplay tool: PyShow.

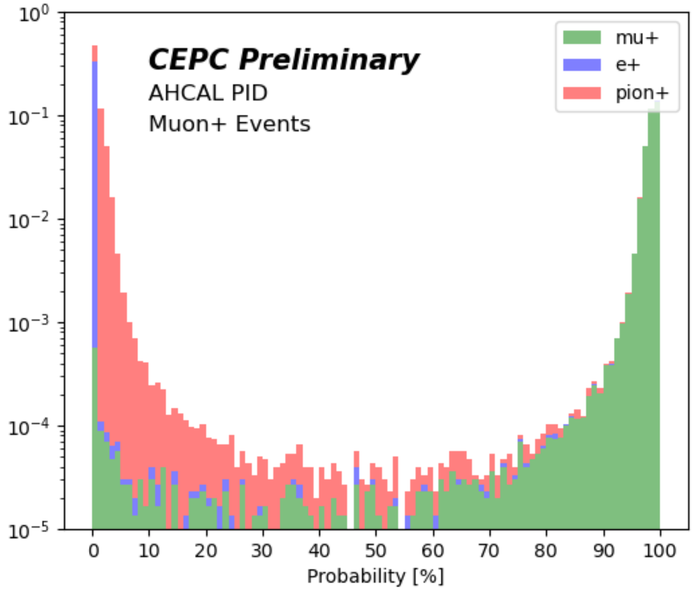
Backup

- Code: https://github.com/Tom0126/CEPC-AHCAL-PID.git

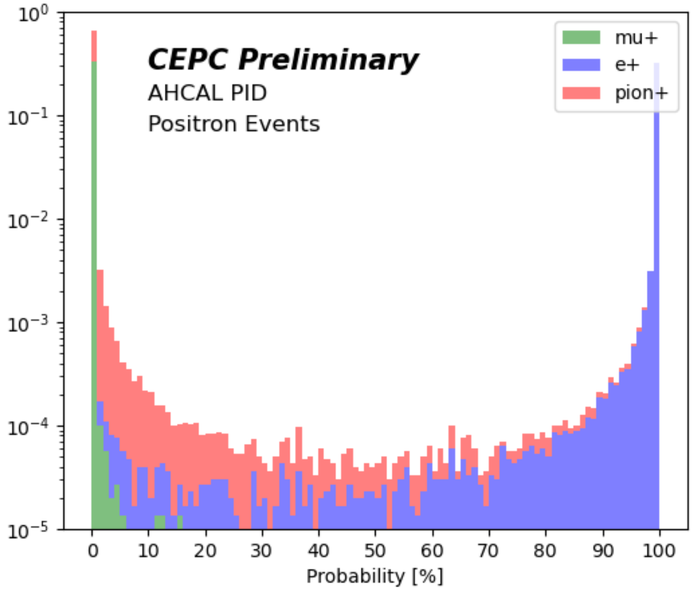


Loss function

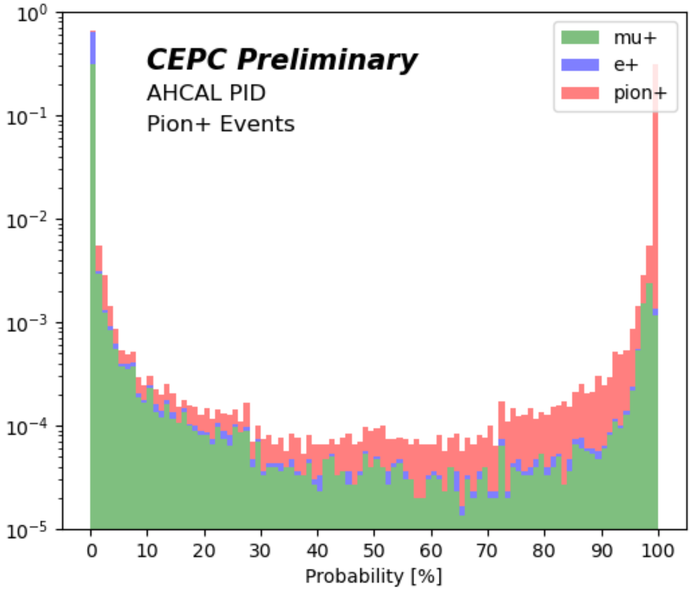
Backup



mu+



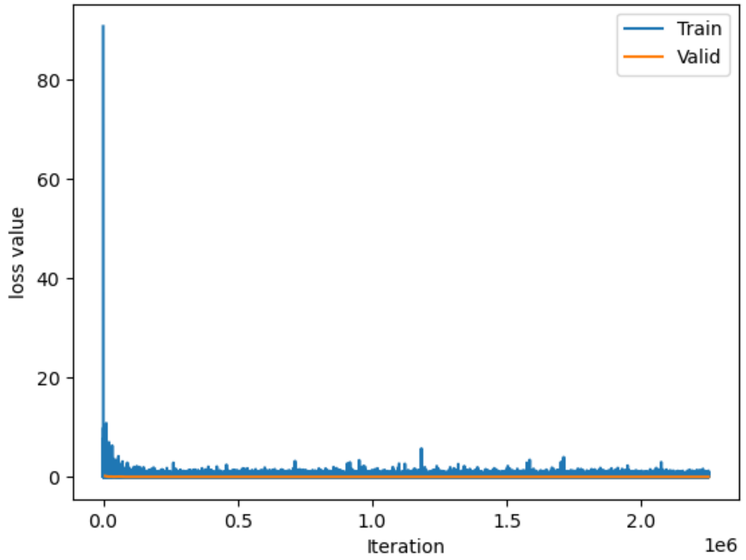
e+



pi+

Log scale

Backup



Loss function