Hyper Kamiokande TRISEP Machine Learning Project

Natnael Debru & Antoine Rehberg & Csaba Velich

Context

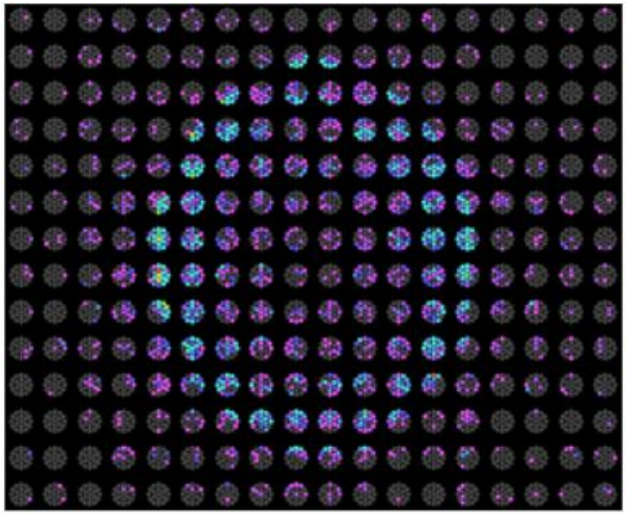
Water Cherenkov Detector

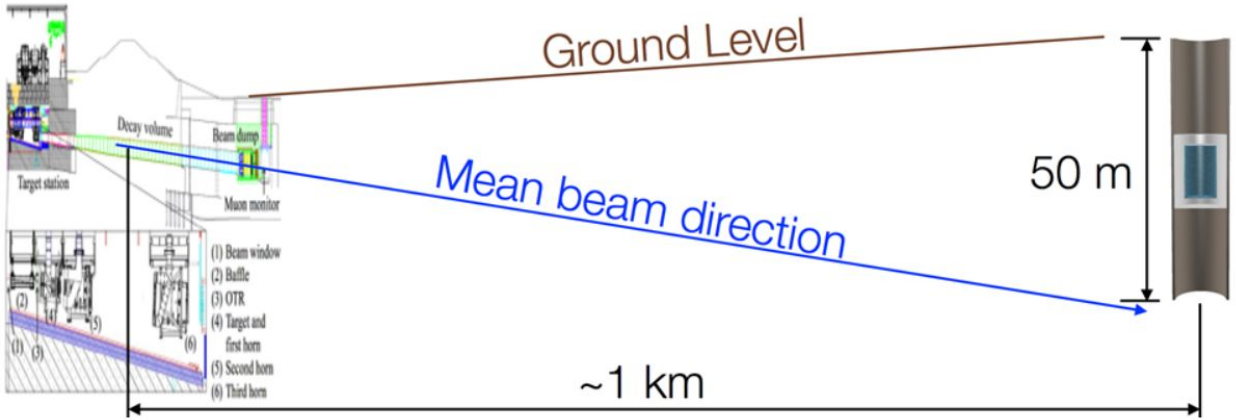
Classify lepton (e / μ) and background (γ)

900k events simulated

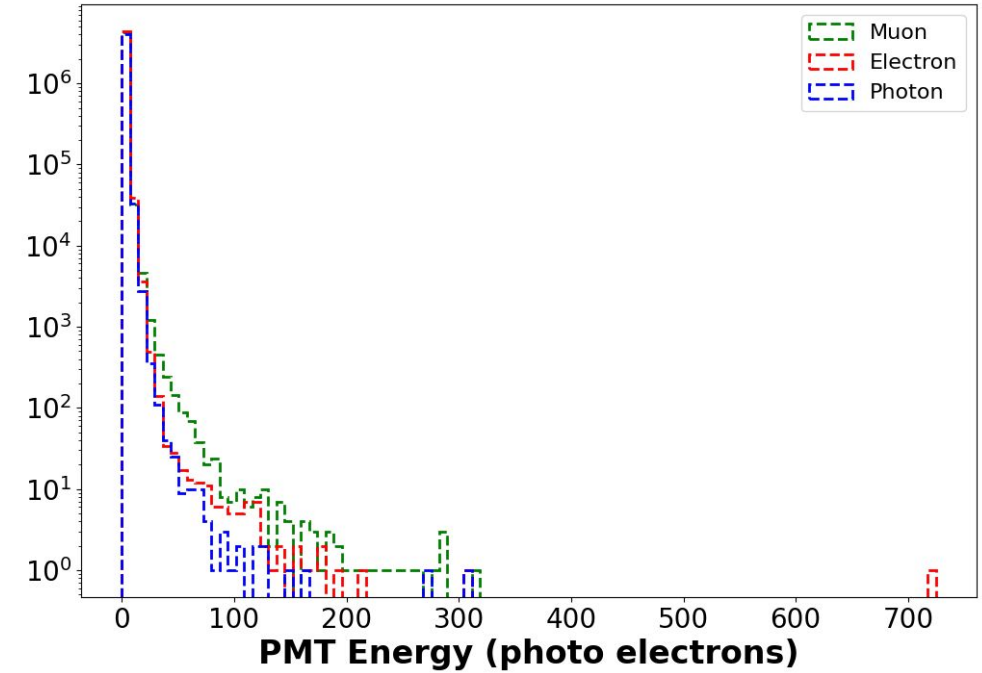
3 labels;

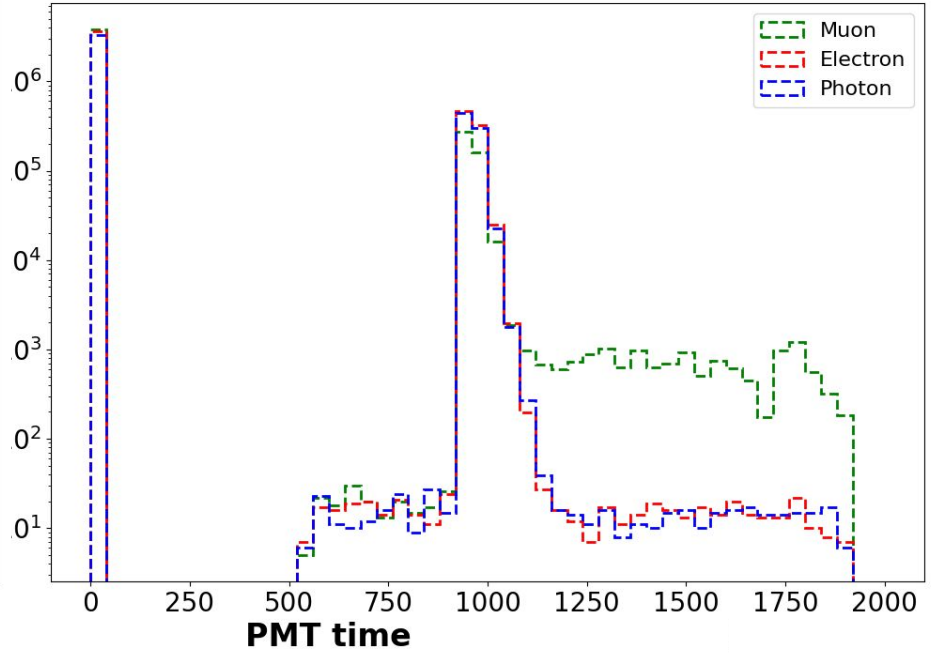
- Label 0 (γ),
- Label 1 (é)
- Label 2 (μ)



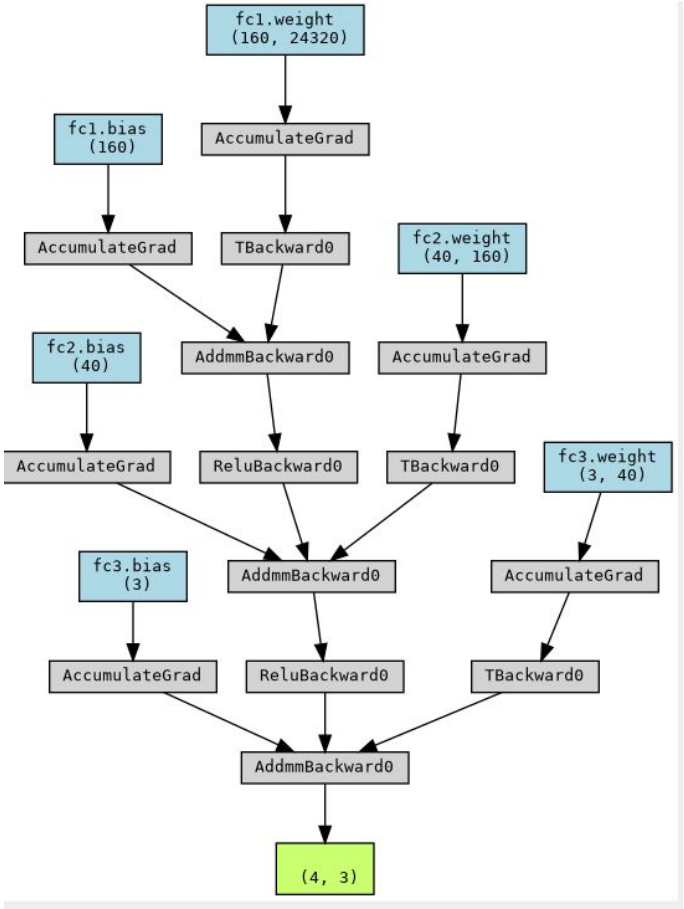


Example of Data





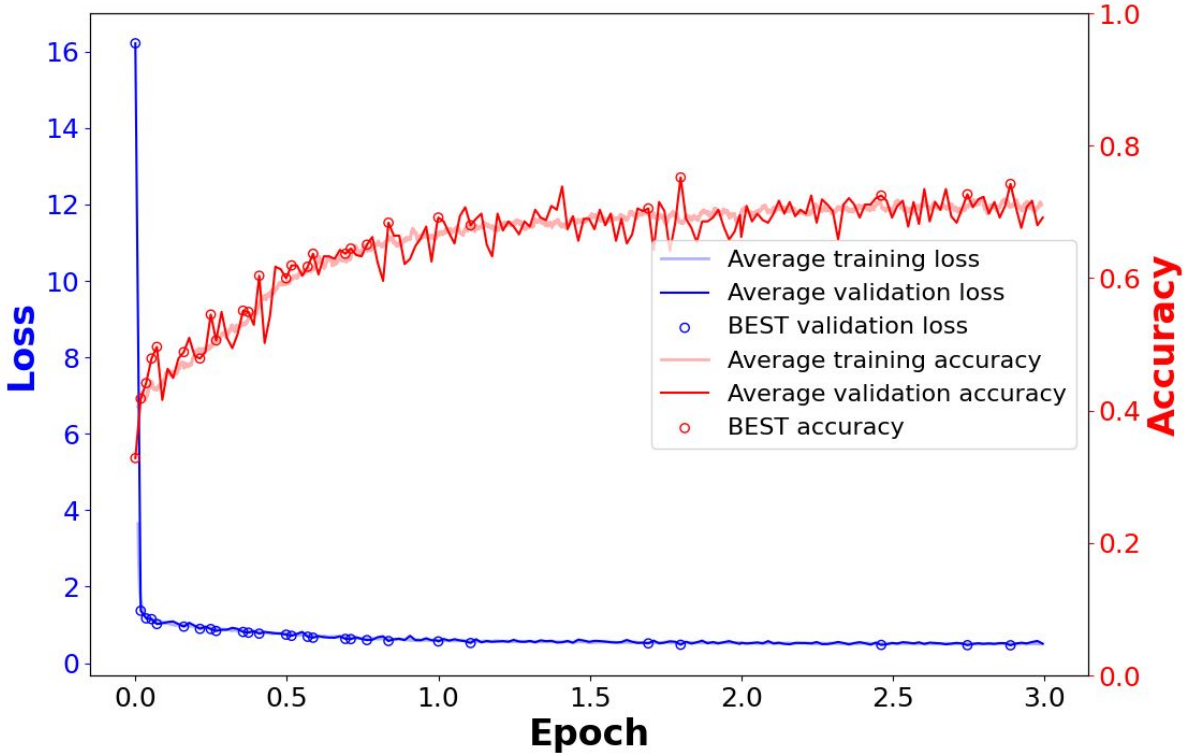
Model 1: MLP



3 fully connected layers
Activation fct: ReLU
Calculate gradient
returns 3 classes

Loss fct:
from torch.nn import
CrossEntropyLoss

Training: MLP



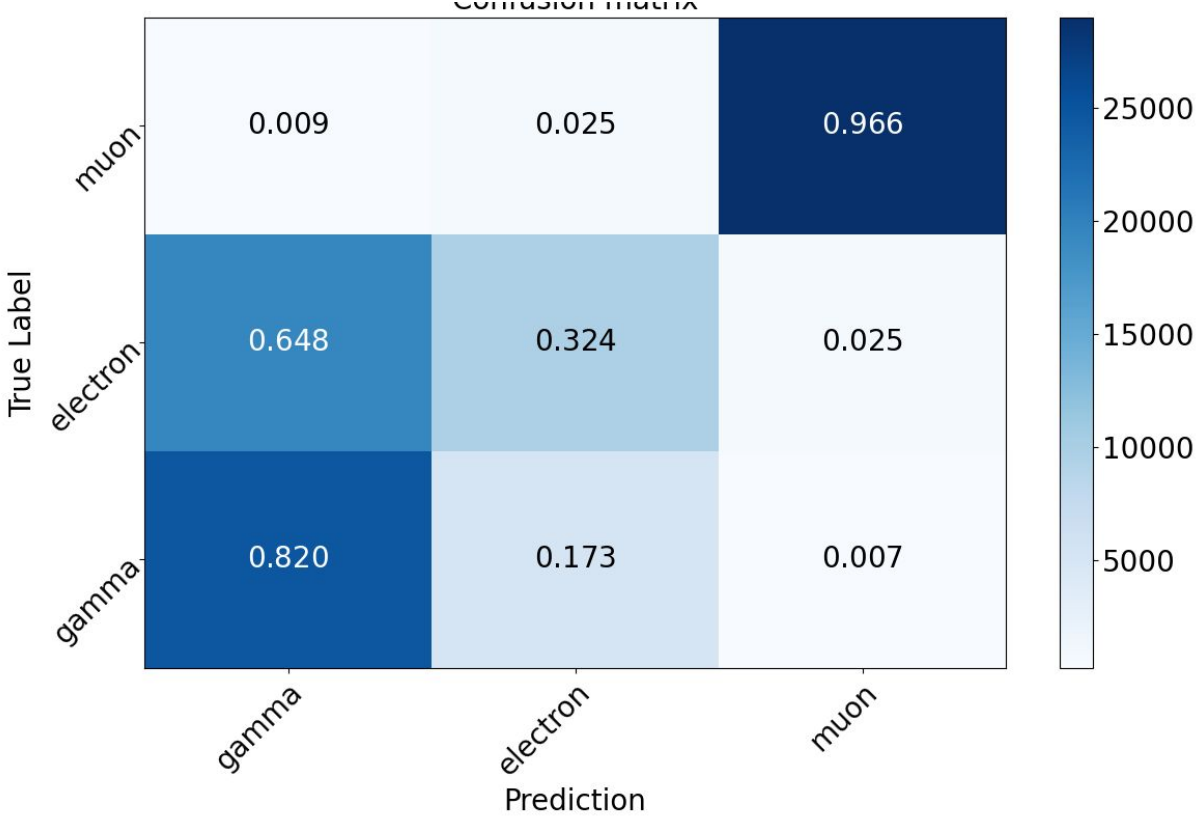
Dataset: total 900k events
80% test
10% validation
10% test

Slow learning rate (0.00001)

3 epochs: loss and accuracy plateau

Model not overtrained or undertrained

MLP Evaluation: Confusion Matrix



Good at identifying muons

~ 65% of electrons mistaken for gamma

~ 20% of gamma mistaken for electrons

Model 2: CNN

Activation function: ReLU

Also outputs 3 classes

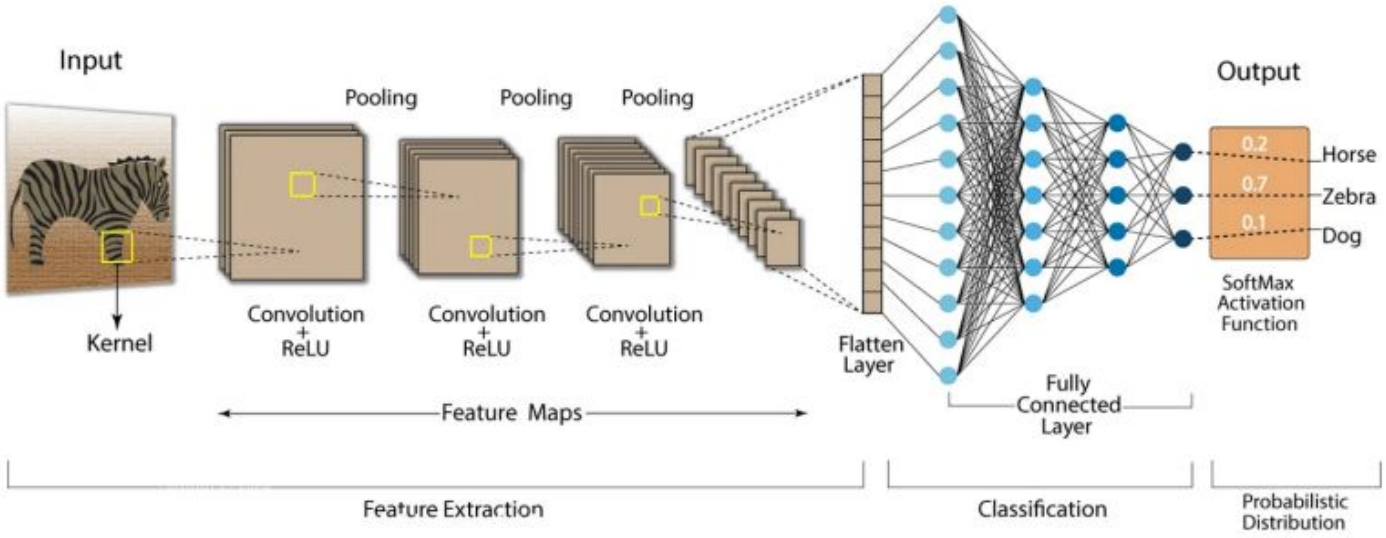
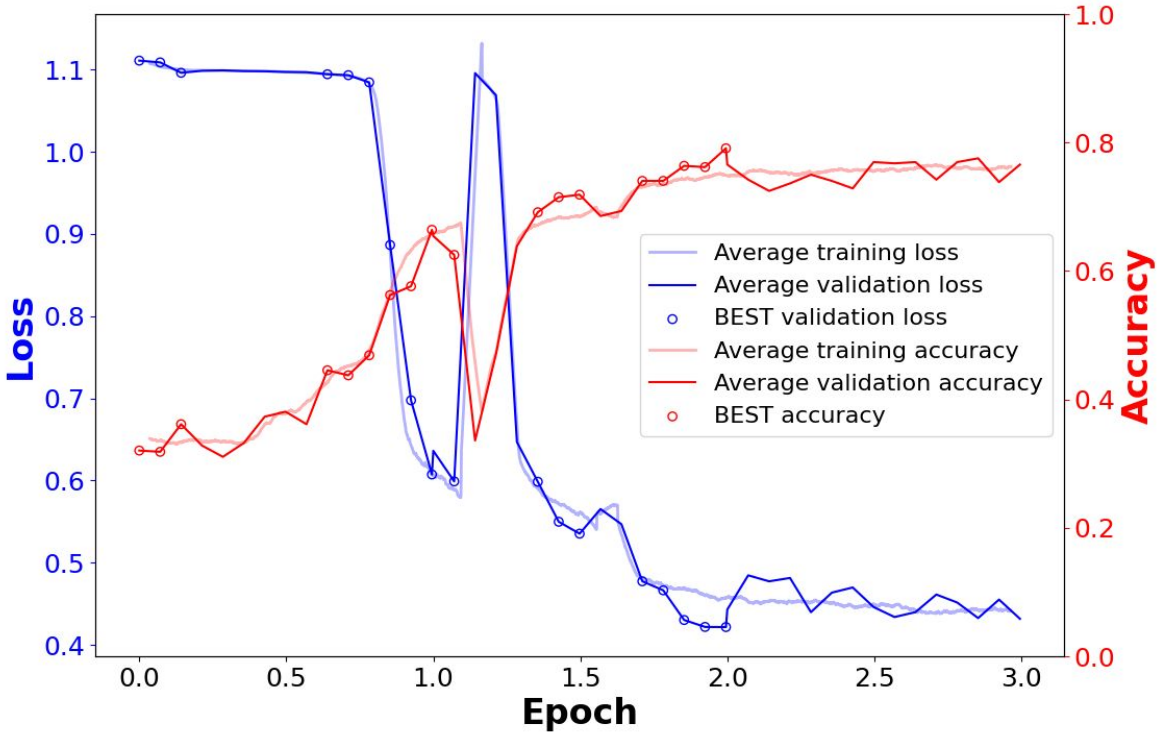


Image stolen from Sean Gasiorowski's slide

Training: CNN



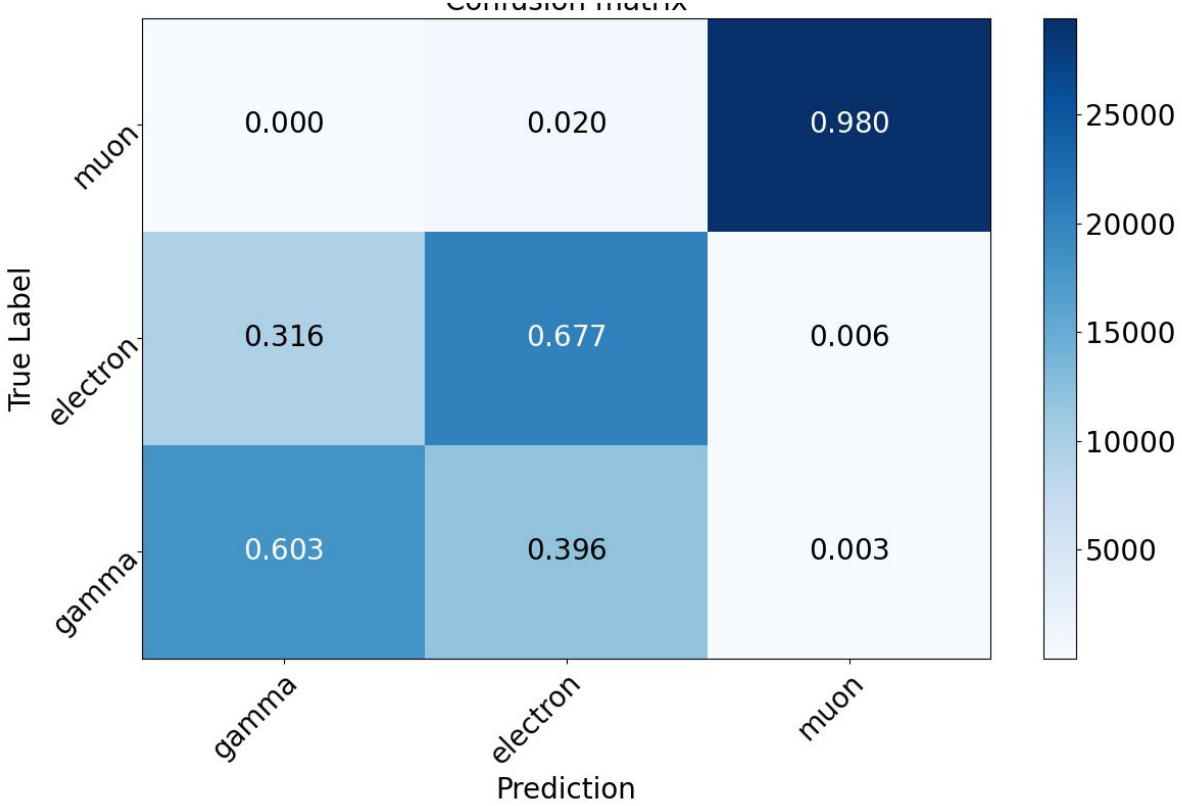
Dataset: total 900k events
80% test
10% validation
10% test

Fast learning rate (0.01)

3 epochs: loss and accuracy plateau

Model not overtrained or undertrained

CNN Evaluation



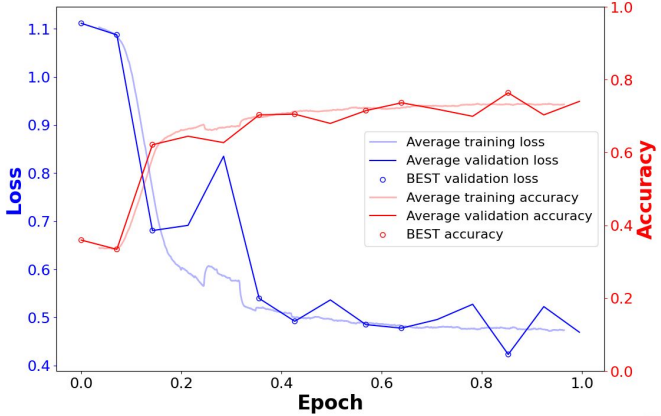
Better at identifying μ

Better at distinguishing electrons and gammas

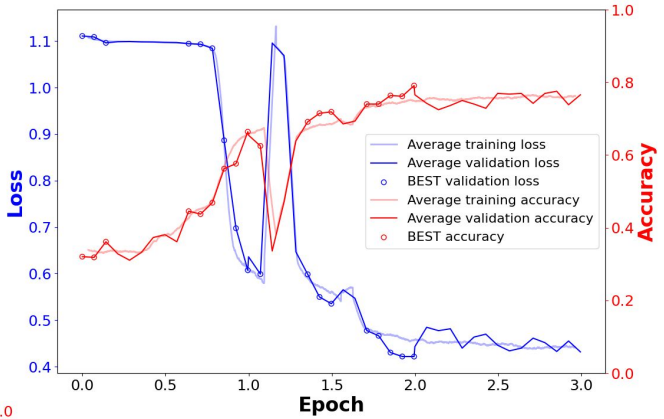
Majority of electrons and gammas (~ 60%) identified correctly

Hyperparameter tuning: Training Time CNN

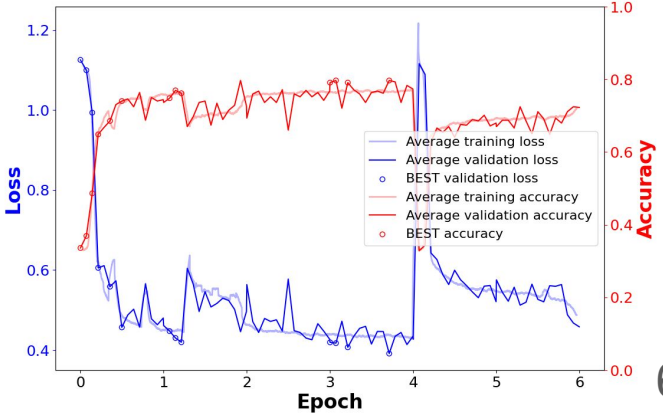
- (Slightly) optimized the training time to avoid undertraining or overtraining



1 epoch



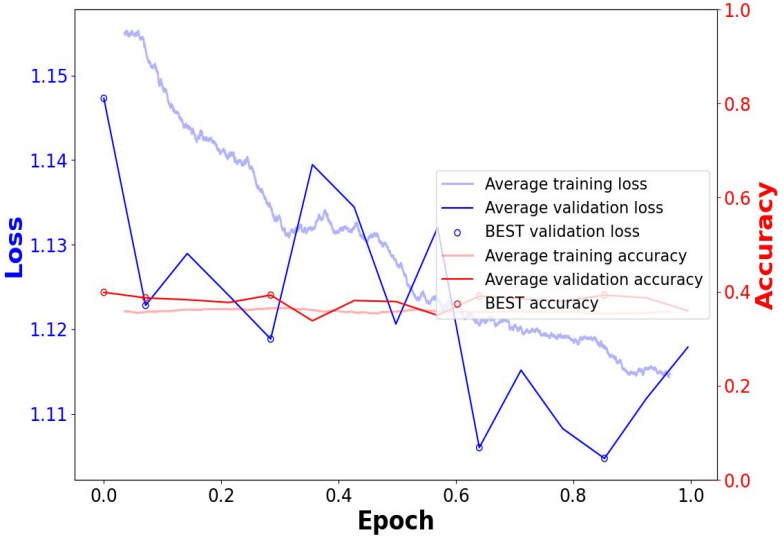
3 epochs



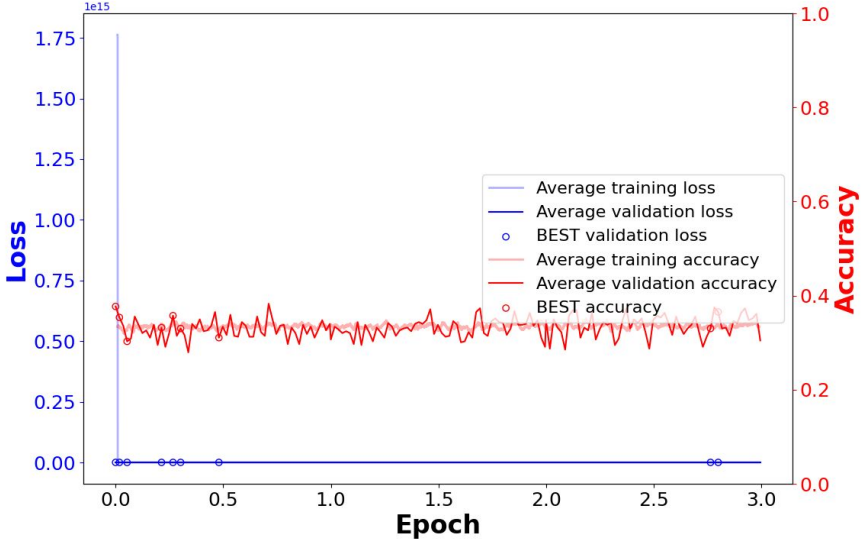
6 epochs

Hyperparameter tuning: Learning rate

CNN with LR=0.00001



MLP with LR=0.01



Different learning rates appropriate for each model

Conclusion

- Use ResNet model
- More fine tuning of hyper-parameters
- Do more pre-processing