KM3NET: MACHINE LEARNING

Chiara De Sio *(Univ. Salerno) for the KM3NeT Collaboration*

cdesio@unisa.it

2 - 4 October 2018 Dubna Russia

VLVnT - 2018
Very Large Volume Neutrino Telescopes

(Some) Existing Machine Learning Applications for KM3NeT

- MVA in Point-Source Analysis (Credits A. Trovato)
  - Random Forest for 3-class Prediction (Source, Atm ν, Atm μ)
- High Energy Starting Muons (Credits K. Pikounis)
  - Boosted Decision Trees for 2-class Prediction (Signal, Background)
- EReNN: Energy Reconstruction with Neural Networks (Credits E. Drakopoulou, et al.)
  - Multi-Layer Perceptron for Energy Reconstruction
- Shallow and Deep Learning Applications in KM3NeT (Credits S. Geißelsöder et al.)
  - Multiple applications of Machine and Deep Learning Models in Supervised and Unsupervised Learning settings

- **Deep Learning applications for ARCA (C. De Sio - me)**
- **Deep Learning in ORCA with OrcaNet (M. Moser et al. - ECAP)**

DESCRIBED IN THIS PRESENTATION

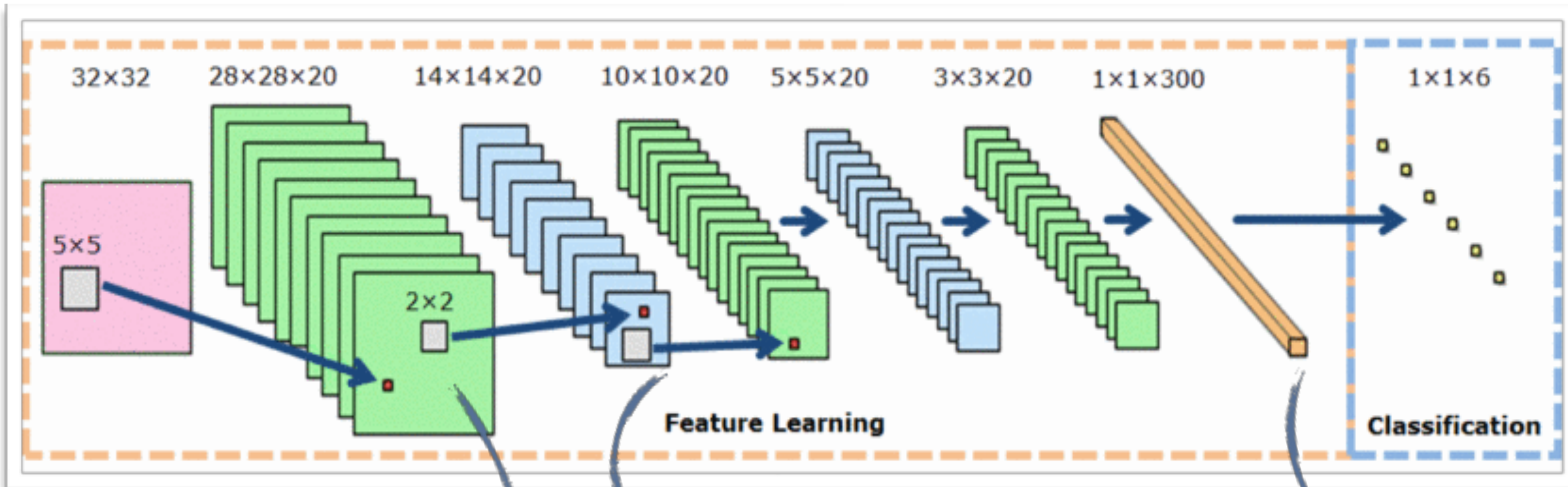
Deep Learning Applications for KM3NeT-ARCA

Four learning tasks:

1) Up-going/Down-going particle Classification
2) $\nu_\mu CC \,/\, \nu_e CC$ interaction Classification
3) Particle Energy Estimation
4) Particle Direction Estimation (Z component)

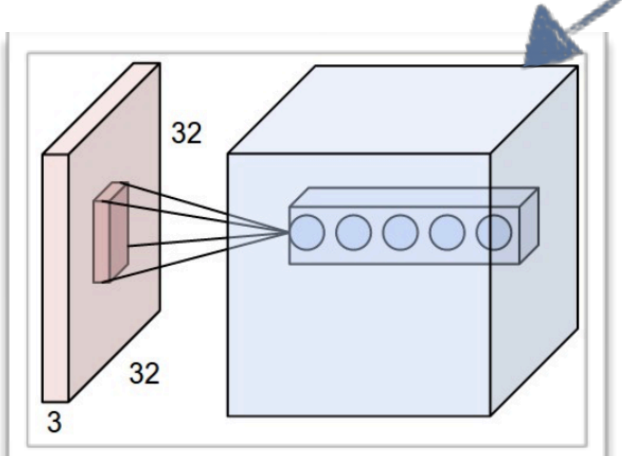
- Using **triggered hits** and times as input data
- Convolutional Neural Network models have been **designed** for each task

Convolutional Deep Networks (CNN): Main Ingredients



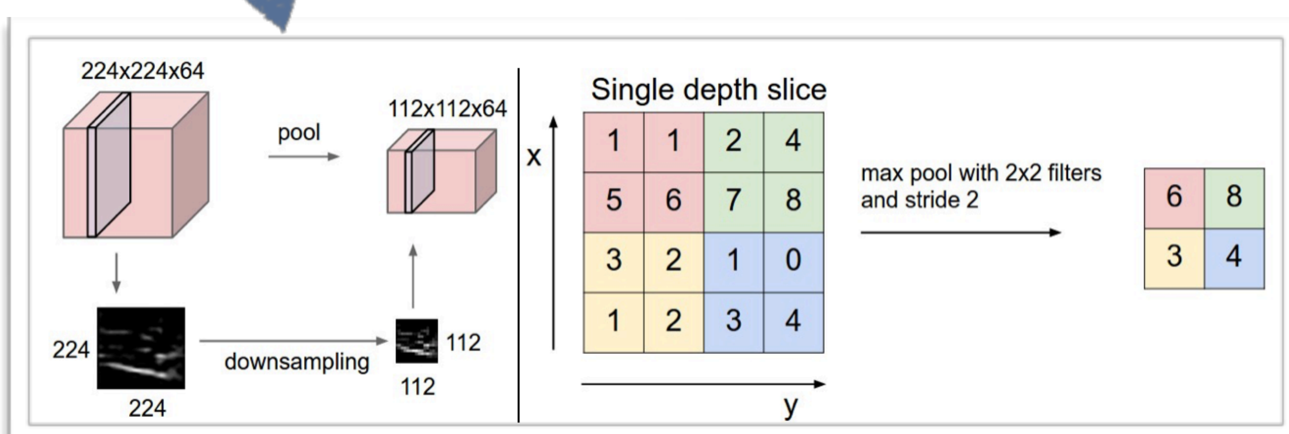
source: bit.ly/cnn-ingredients

Convolutional Layer



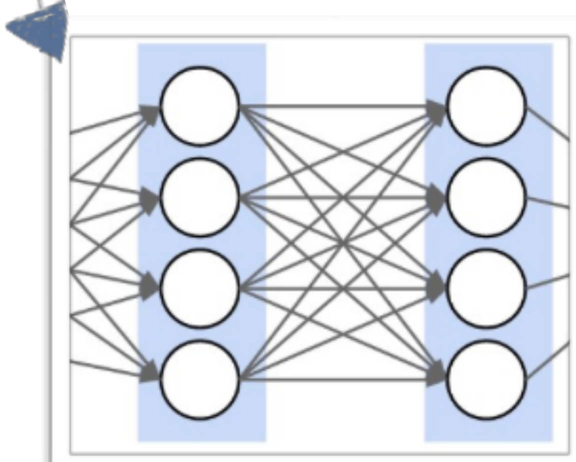
Local Feature Maps Learning

Pooling Layer



Downscaling and Space Invariant Features Learning

Dense Layer

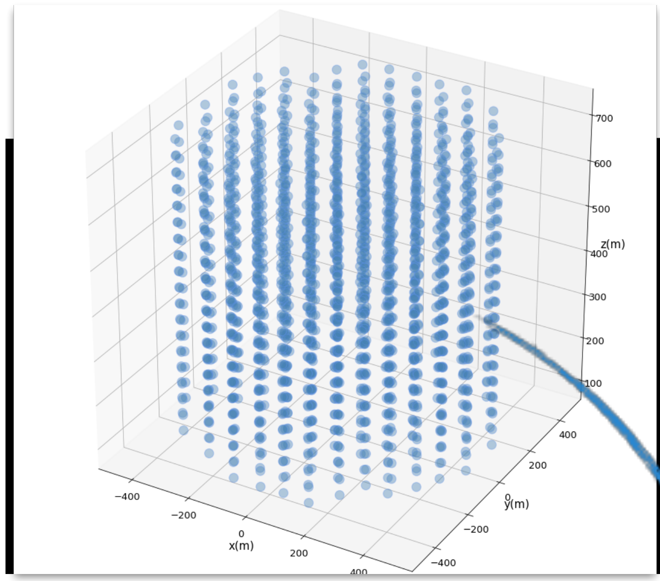


Global Feature Learning & Prediction

DATA PREPARATION

PREPARE DATA TO BE FED INTO NEURAL NETWORKS

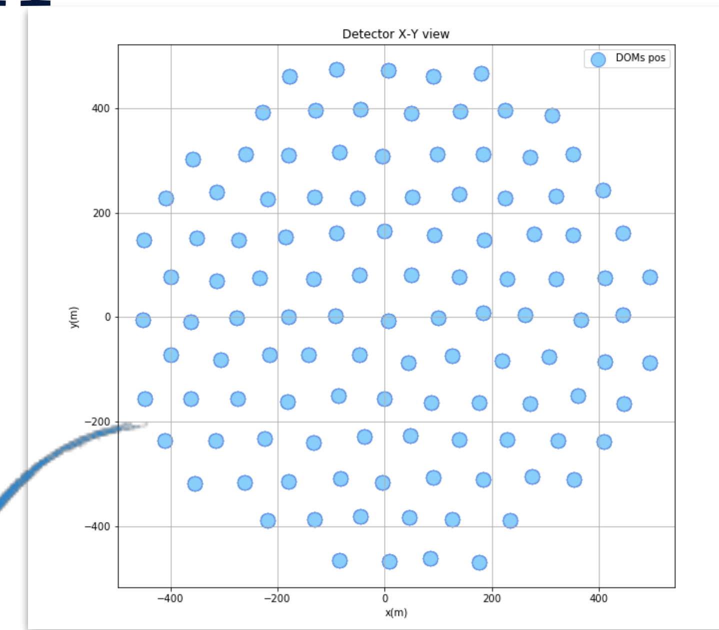
Space regularisation

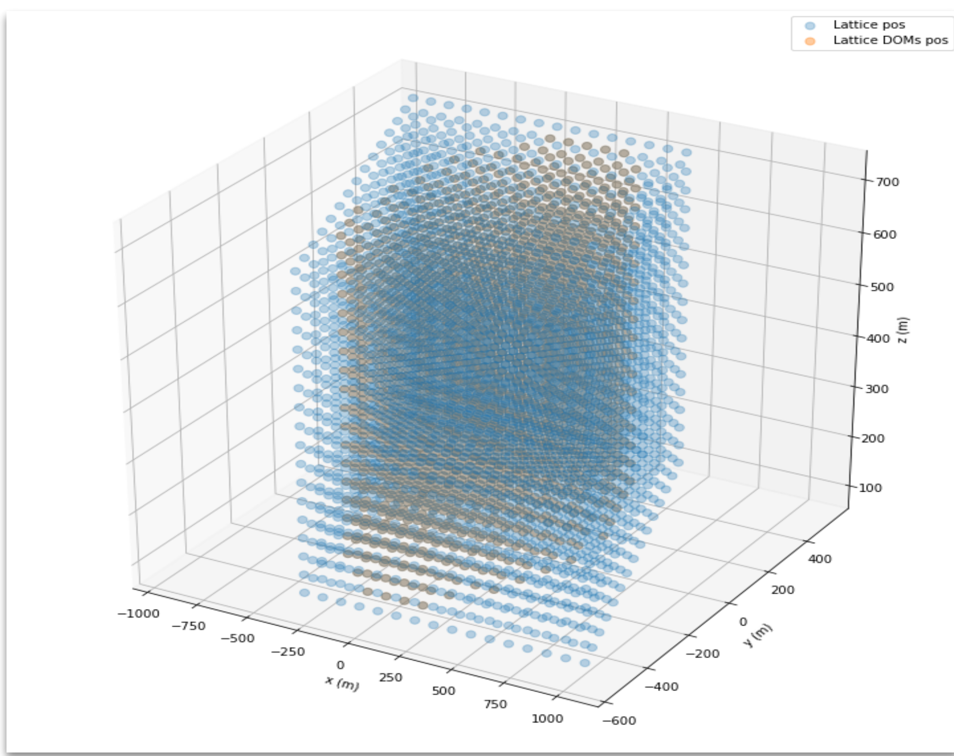


Regularised Detector Structure*

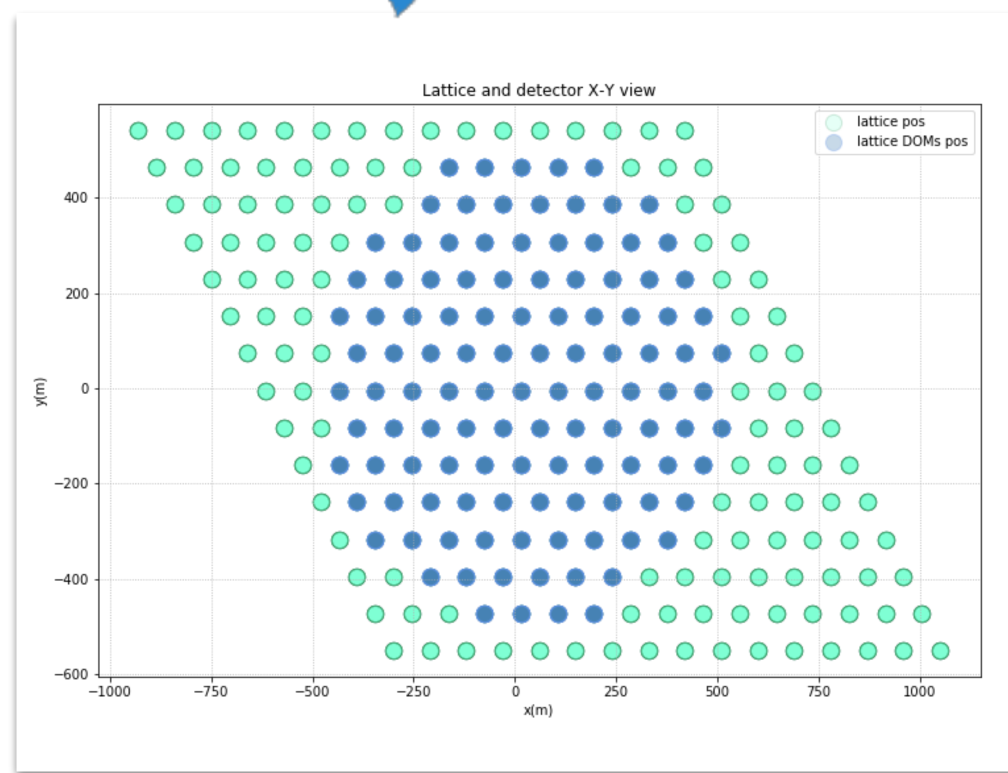
- **exactly** 90m spaced in (X,Y)
- **exactly** 36m spaced in Z

Regularised detector contained in **Lattice**





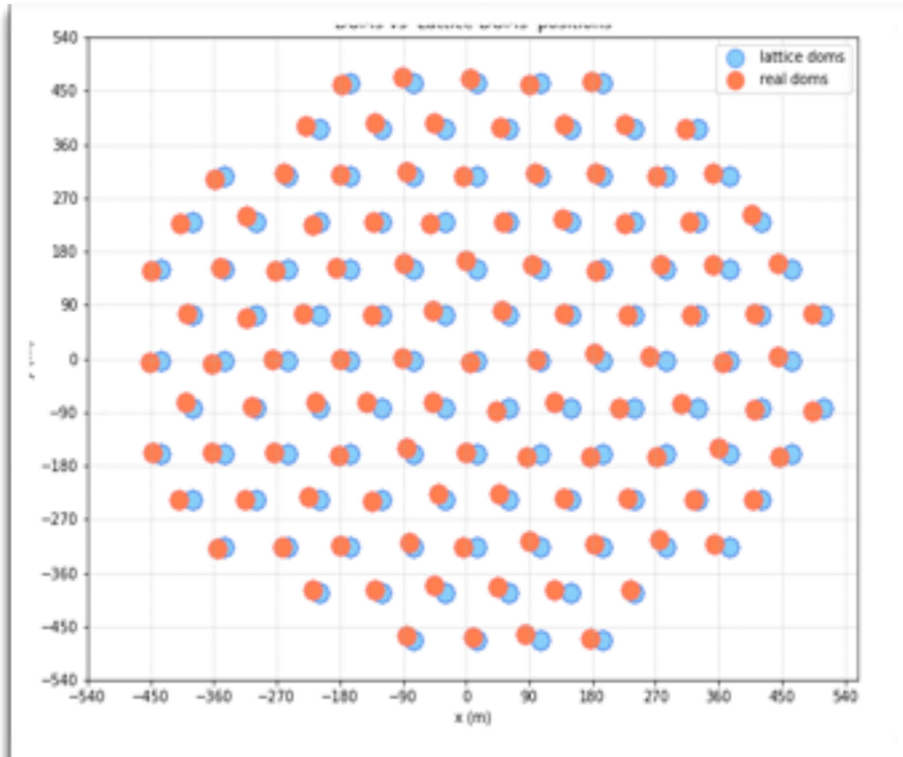
Regularised Detector XYZ-view



Regularised Detector XY-view

*Deviation from regularised structure can be introduced later as a next-order correction

Event Definition



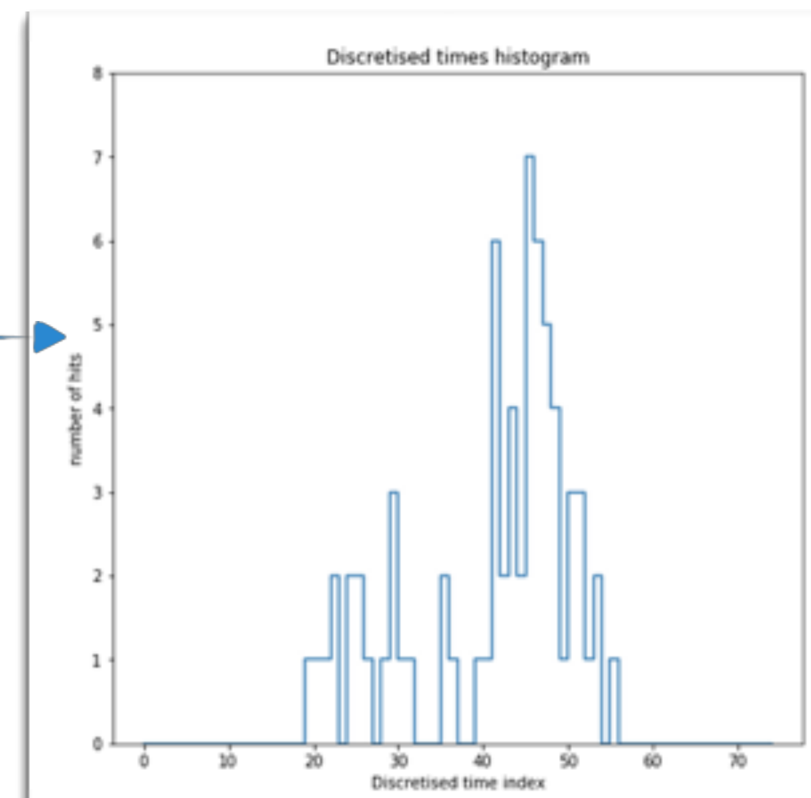
Lattice DOMs vs Real DOMs Positions

Reducing (useless and time consuming) sparsity in data

- Transforming DOM IDs into Lattice DOM IDs
  - a single DOM ID is mapped to an index in the [16x15x18] Lattice

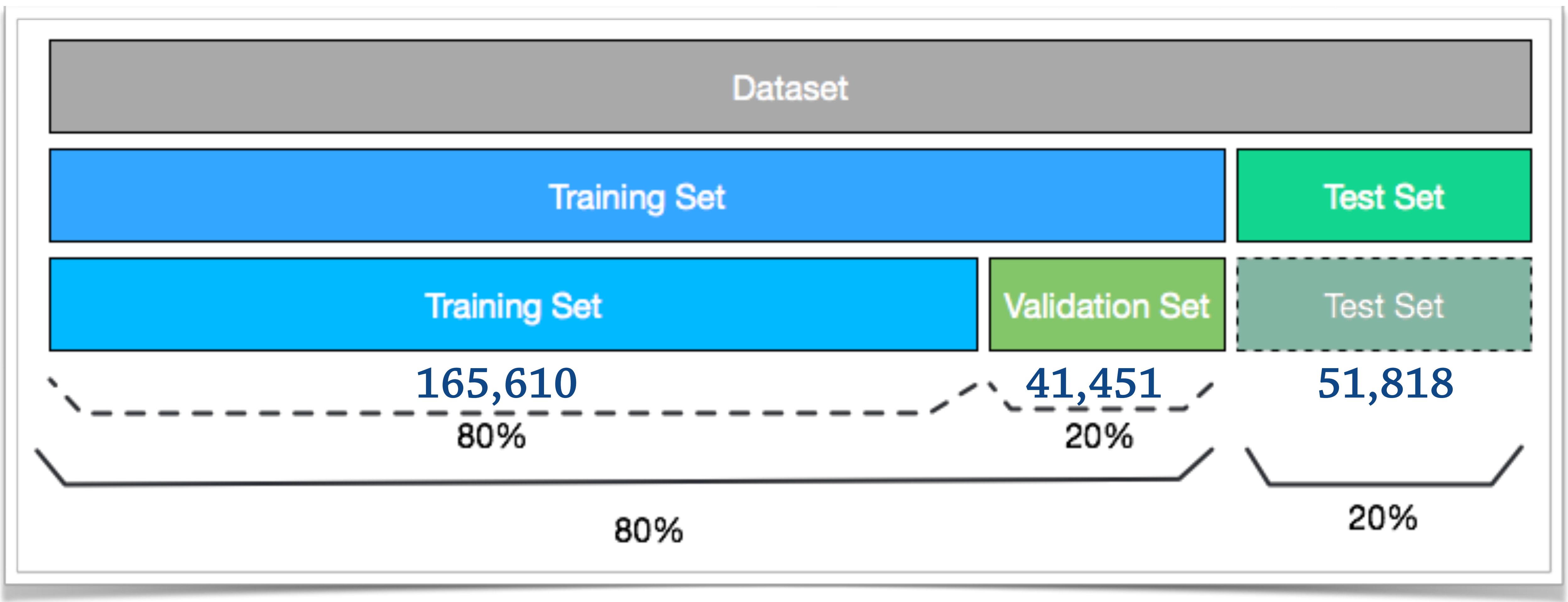
Event in the Regularised Detector:

1. Times are discretized, with a fixed number of steps (i.e. n_Time_steps=75) ~12 ns/bin
2. A collection of **hits** over time
   - Event mapped to a 4D structure (tensor) of shape [75x16x15x18]
   - [discrete_time_index, x_index, y_index, z_index]



Dataset

258,879 total events (samples) arranged from 100 $\nu_e CC$ + 100 $\nu_\mu CC$ files



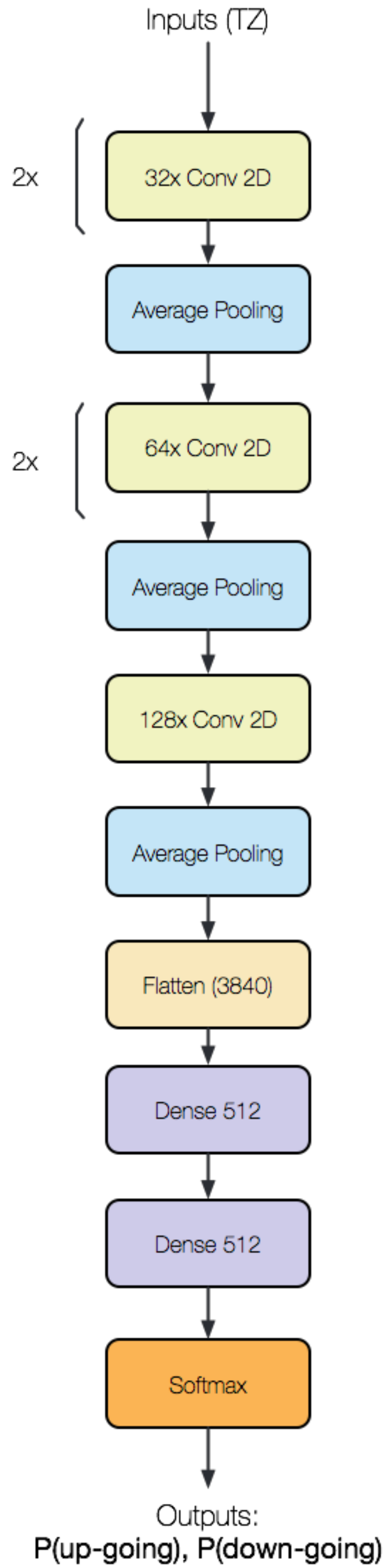
GEFORCE GTX 1080

scikit learn + TensorFlow + Keras

LEARNING TASK 1. UP-GOING/DOWN-GOING NEUTRINO CLASSIFICATION

CLASSIFY UP-GOING AND DOWN-GOING NEUTRINOS ACCORDING TO THEIR Z-COORDINATE EVOLUTION OVER TIME

Model architecture and Input Data



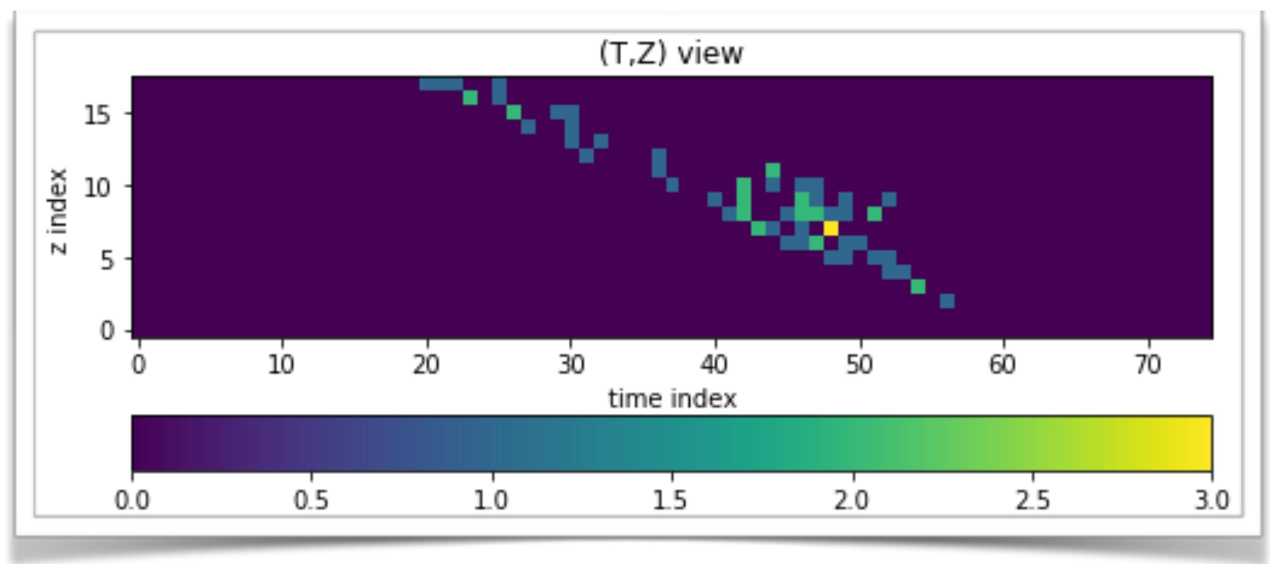
Labels

$$y: \begin{cases} \cos(\theta_z) > 0 : \text{“up-going"} \\ \cos(\theta_z) \leq 0 : \text{“down-going”} \end{cases}$$

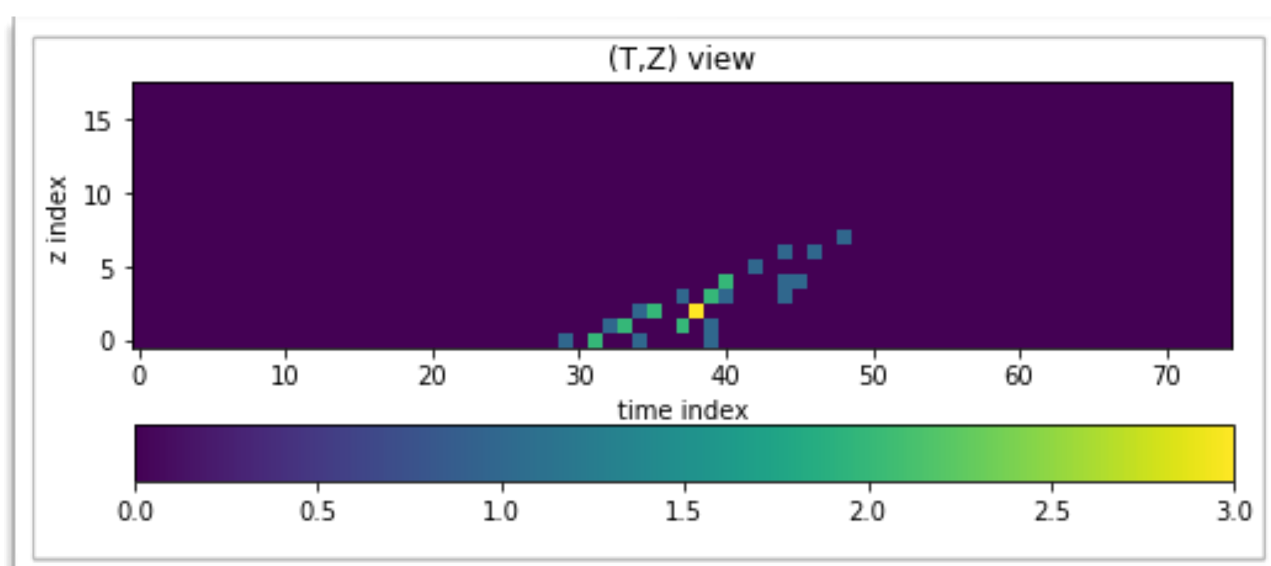
Input tensors reshaped to:

(T,Z): [n_samples, 75, 18]

Input array summed over X and Y axes



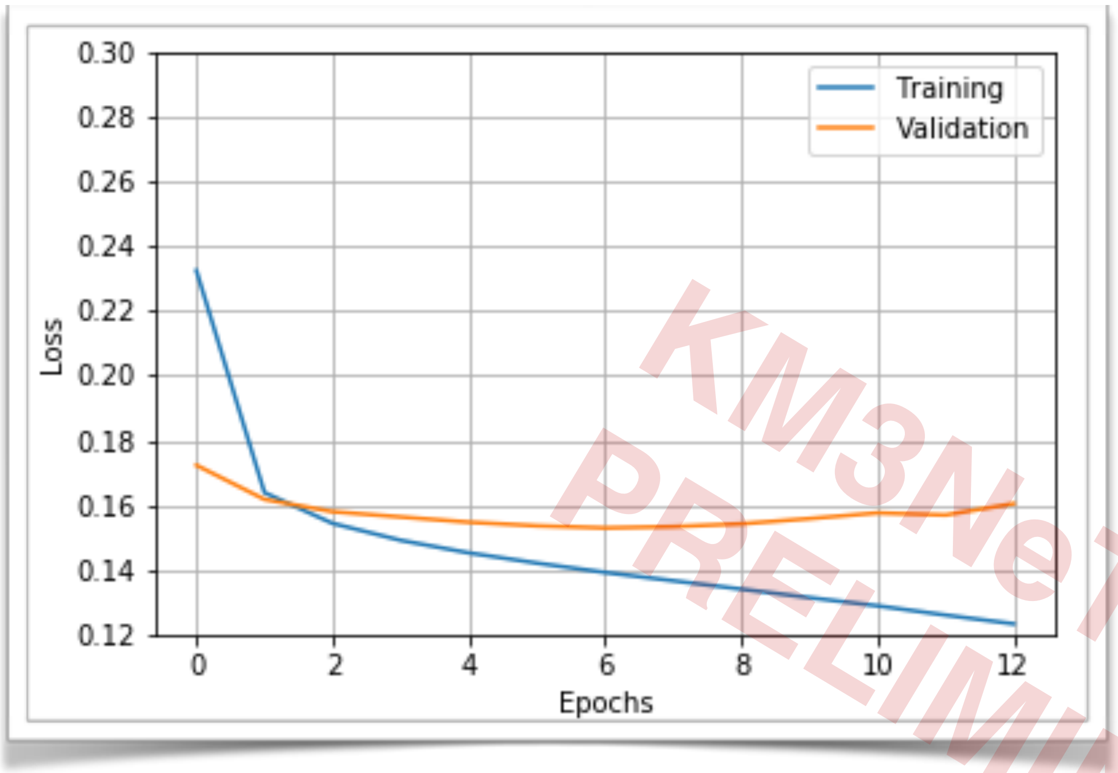
Down-going

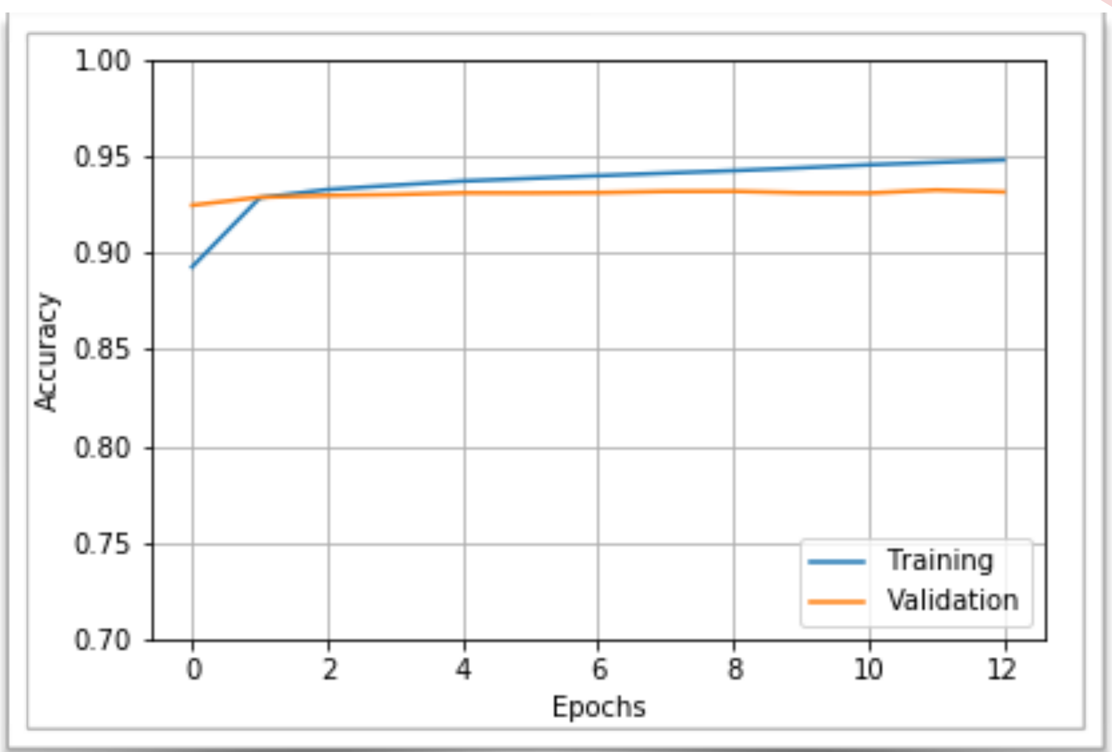


Up-going

Classification results

Training history

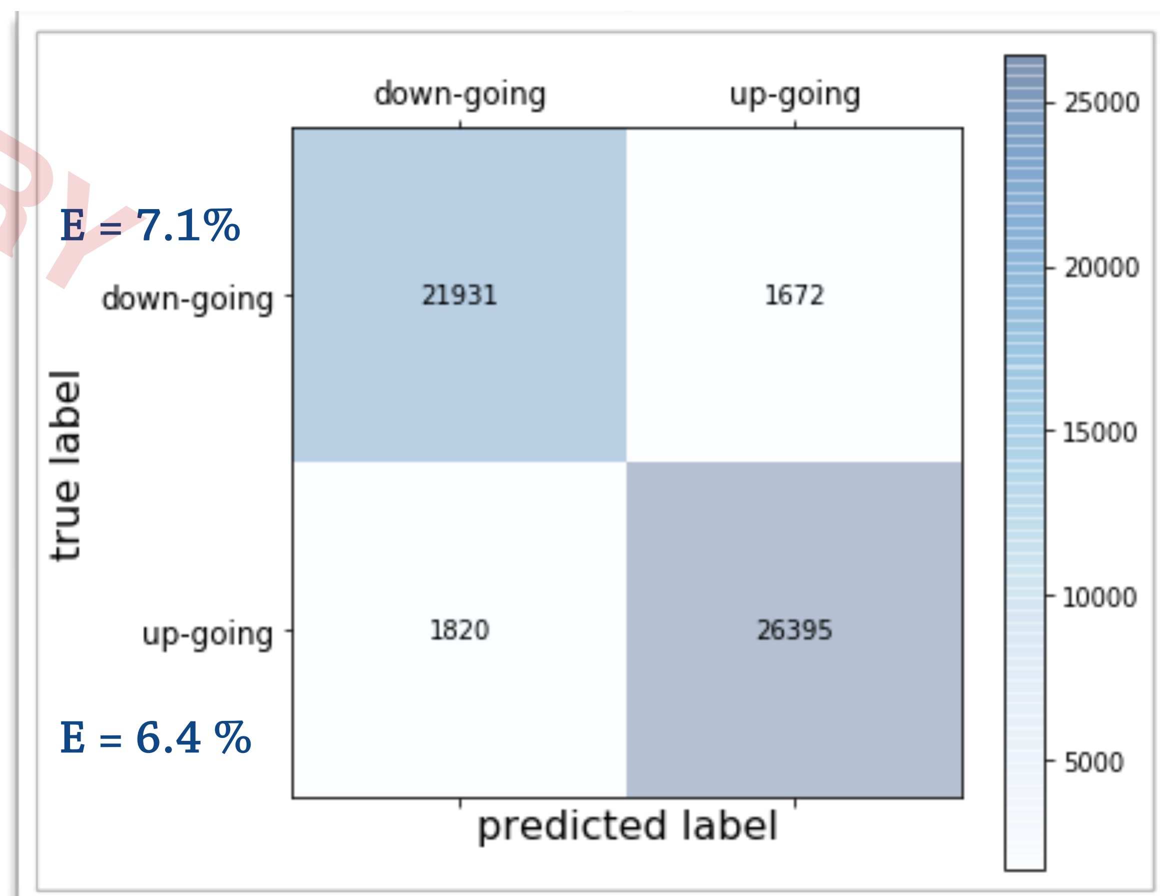




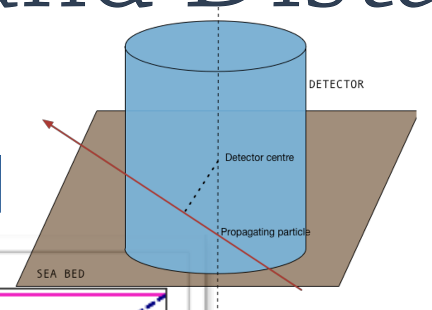
Classification accuracy (on test data)

93.3 %

Confusion matrix (on test data)



Energy and Distance dependency

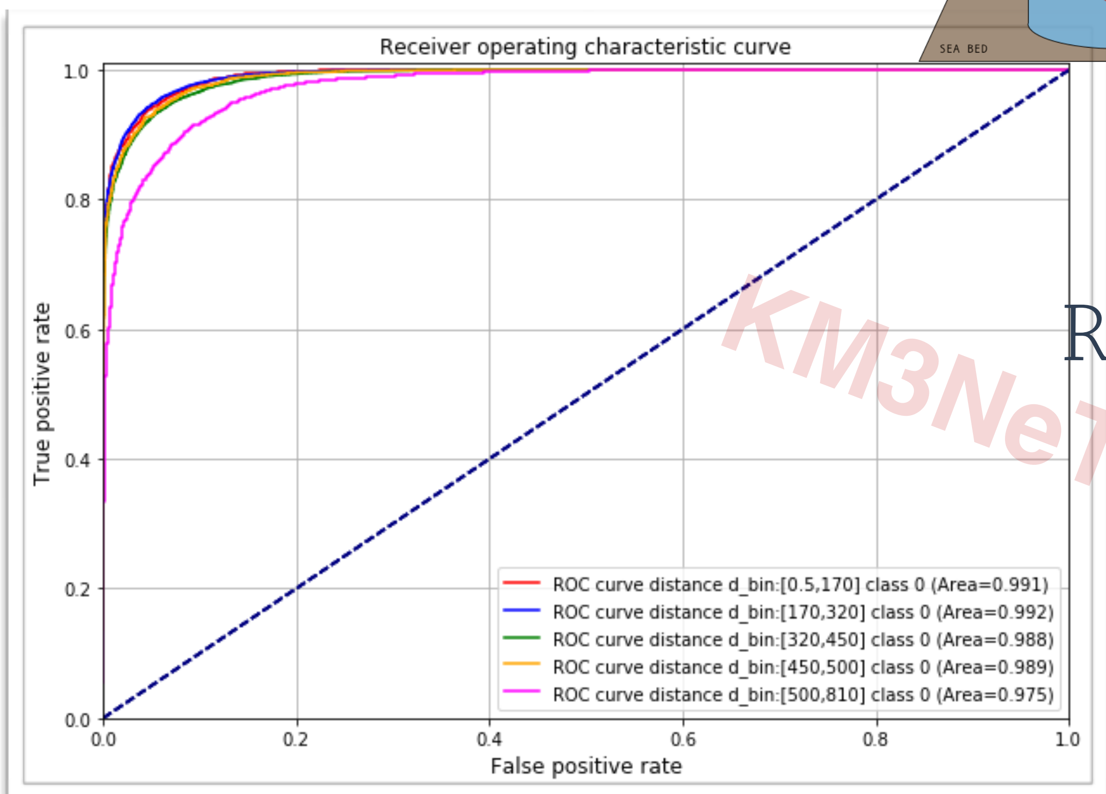


Distance bins (m)

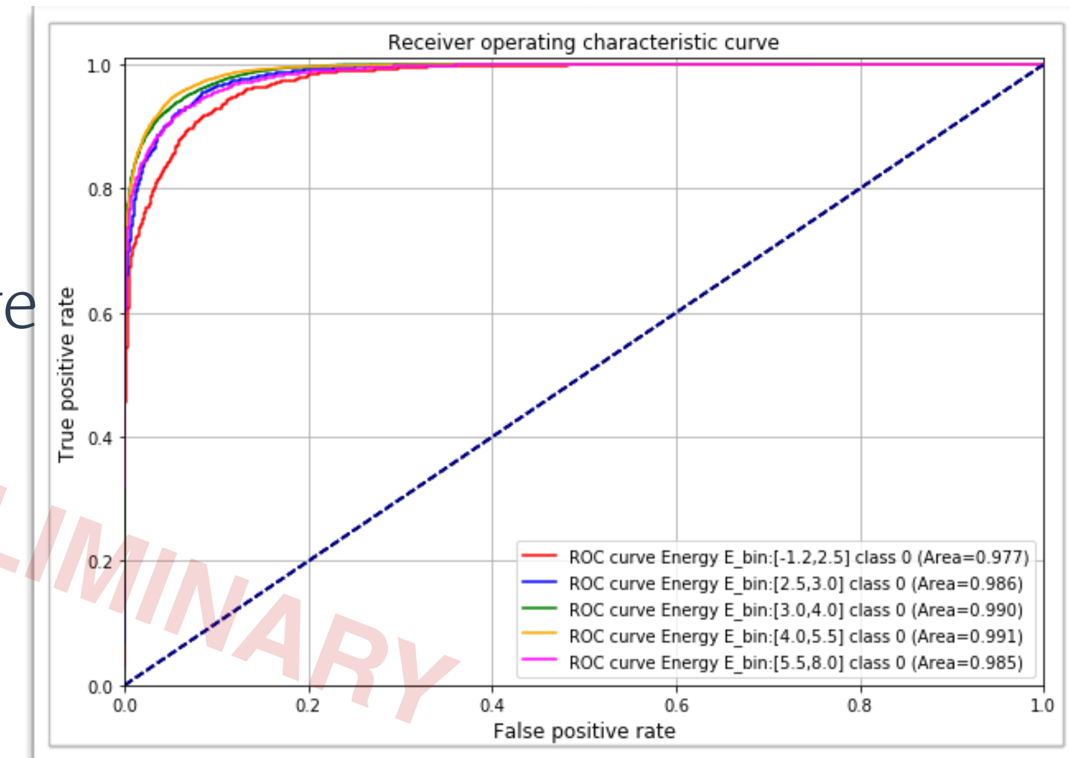
[0.5, 170, 320, 450, 500, 810]

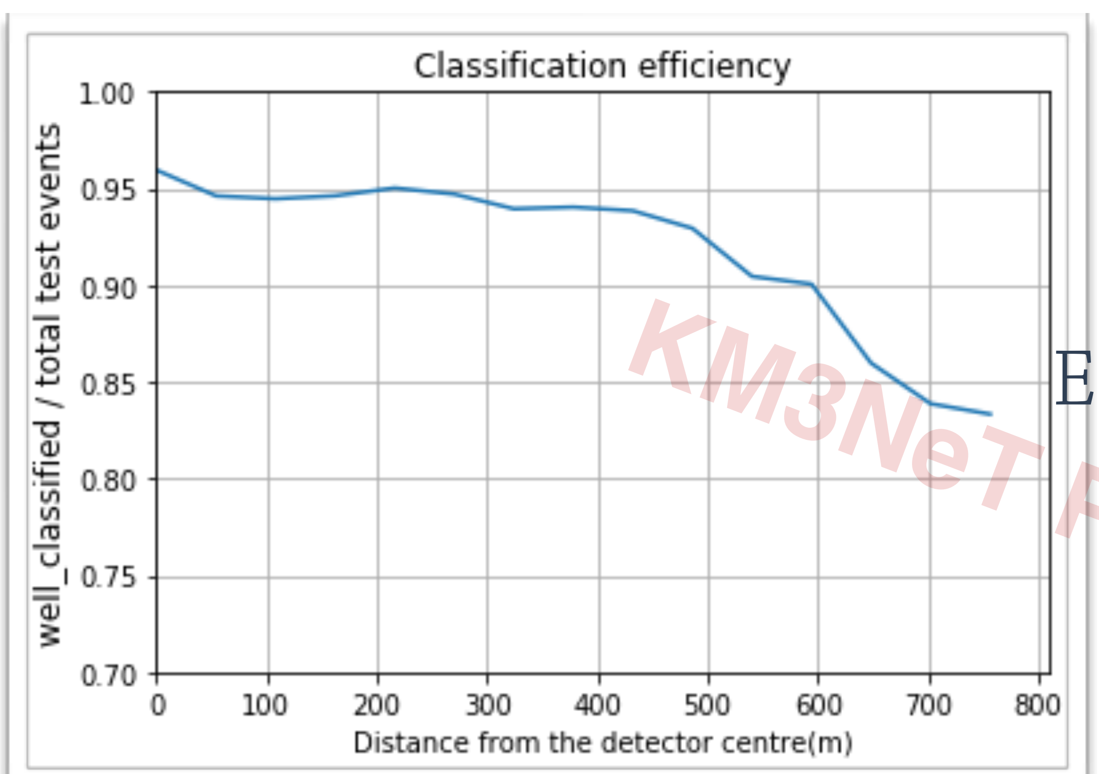
Energy bins (Log(E)) [GeV]

[-1.2, 2.5, 3.0, 4.0,5.5, 8.0]

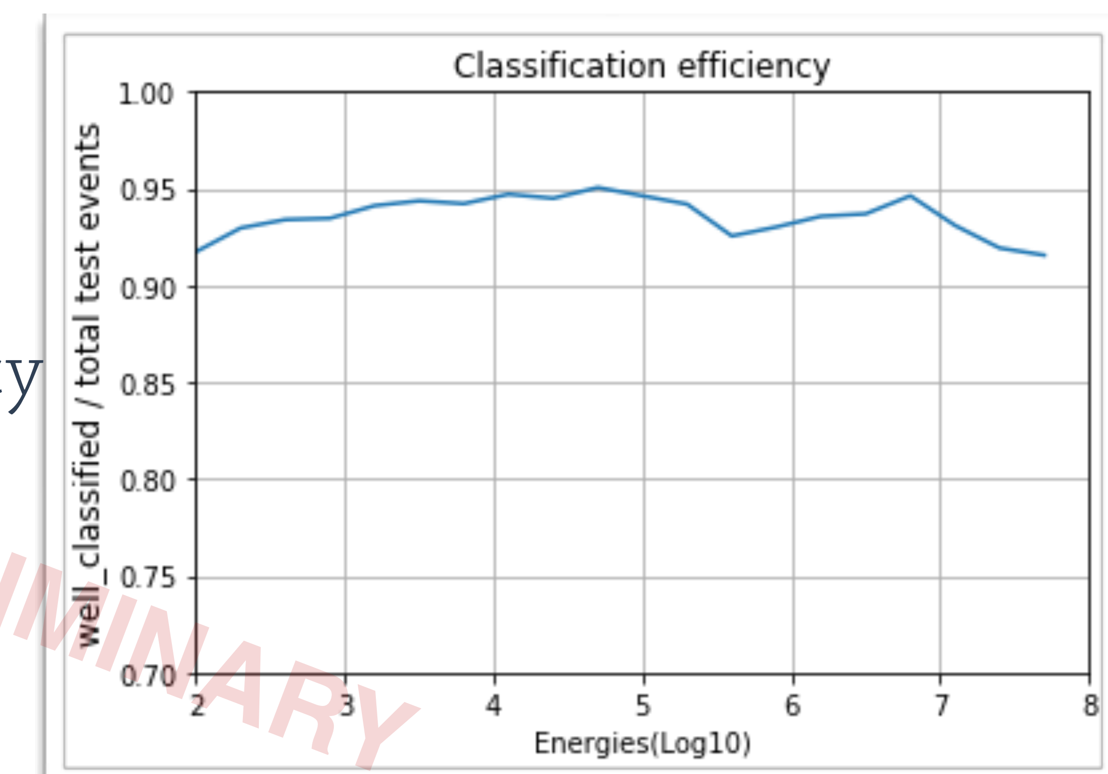


ROC curve





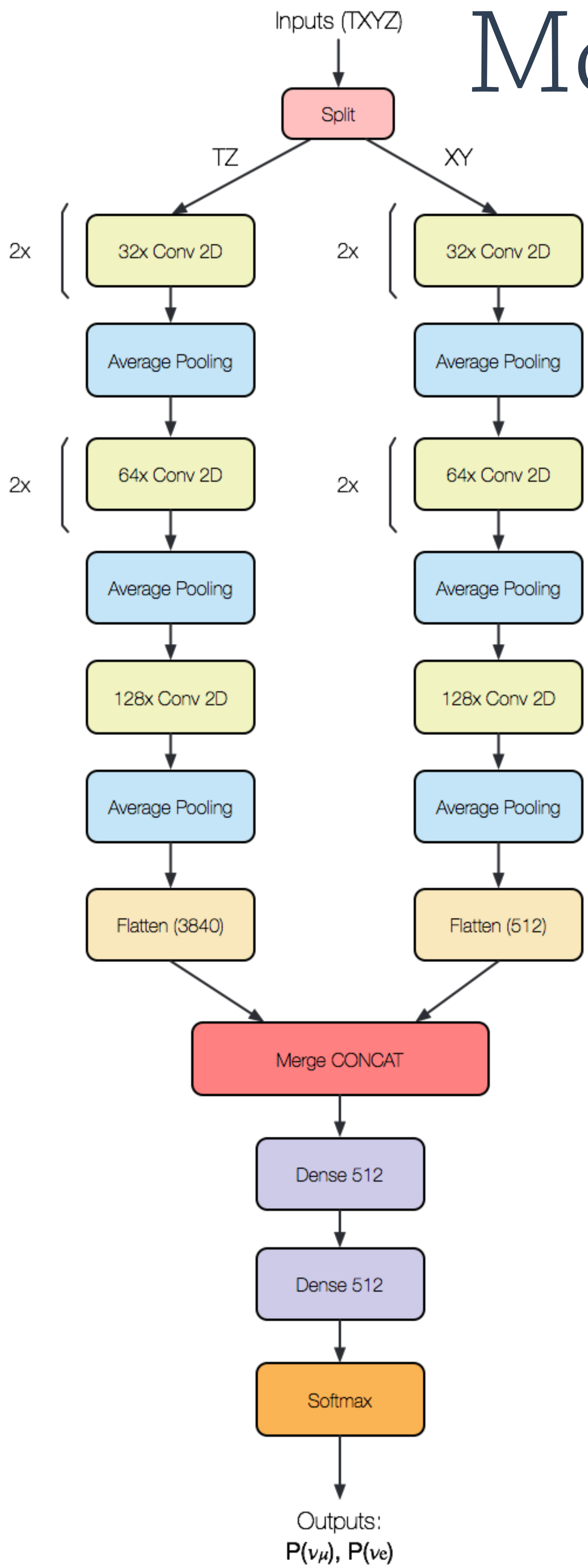
Efficiency



LEARNING TASK 2. $\nu_\mu \mathrm{CC} / \nu_\mathrm{e} \mathrm{CC}$ INTERACTION CLASSIFICATION

CLASSIFY $\nu_\mu CC$ / $\nu_e CC$ INTERACTIONS BASED ON THE SHAPE OF THE EVENTS AND THE EVOLUTION OF POSITIONS OVER TIME

Model architecture and Input Data



CNN model with parallel branches analysing (T,Z) and (X,Y) evolution separately, merged to extract common features

2 Input tensors of shapes:

(T,Z): [n_samples, discrete_time_index, z_index]

(X,Y): [n_samples, x_index, y_index]

Labels

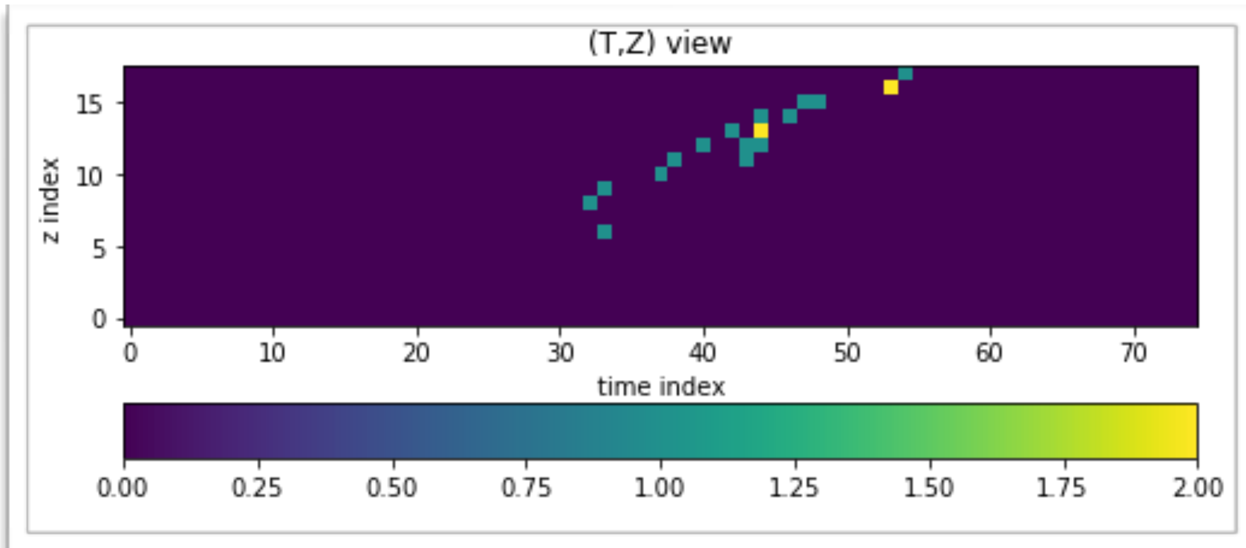
$$y = \begin{cases} 0: & \nu_\mu CC \\ 1: & \nu_e CC \end{cases}$$

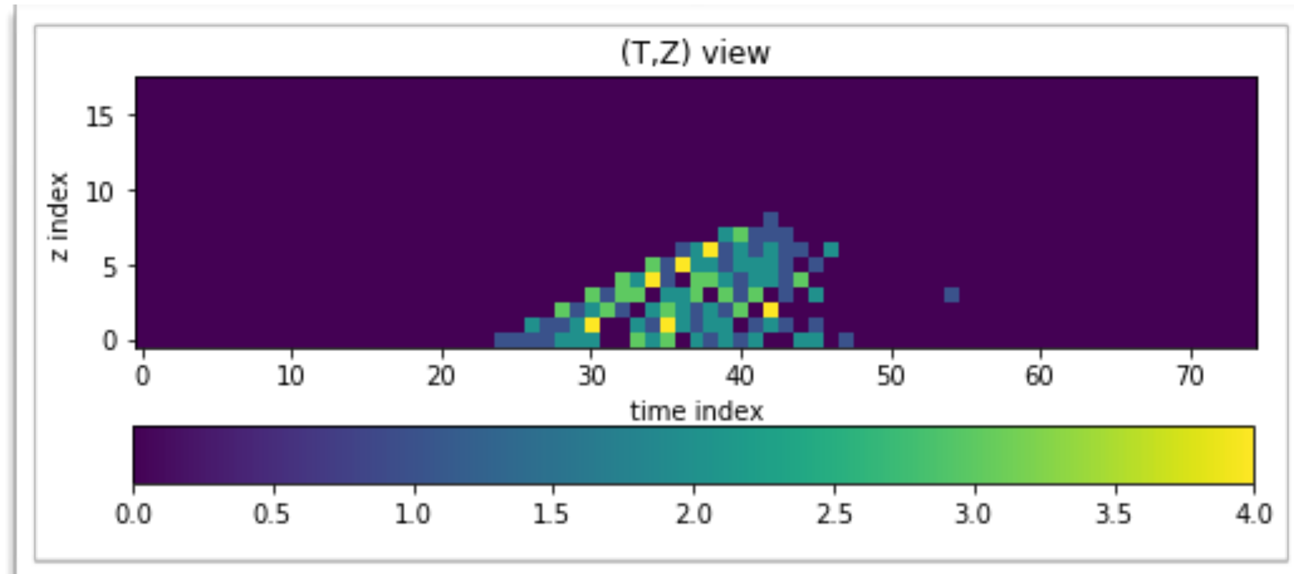
Multiple Views of Events

ν_μCC

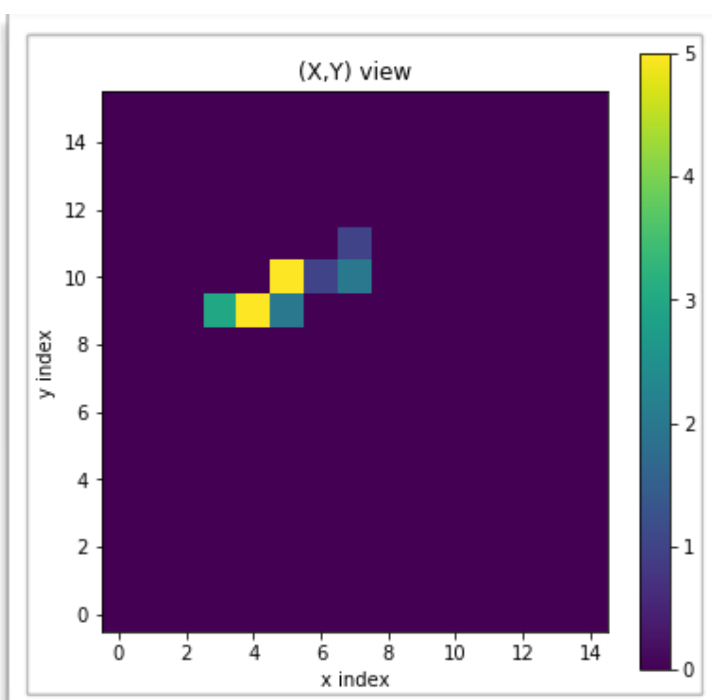
TZ View

ν_eCC

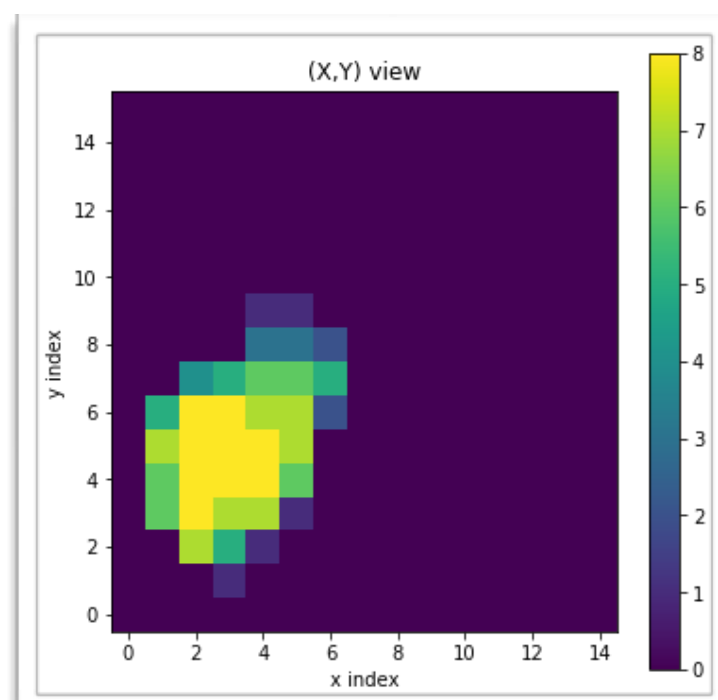




600 GeV



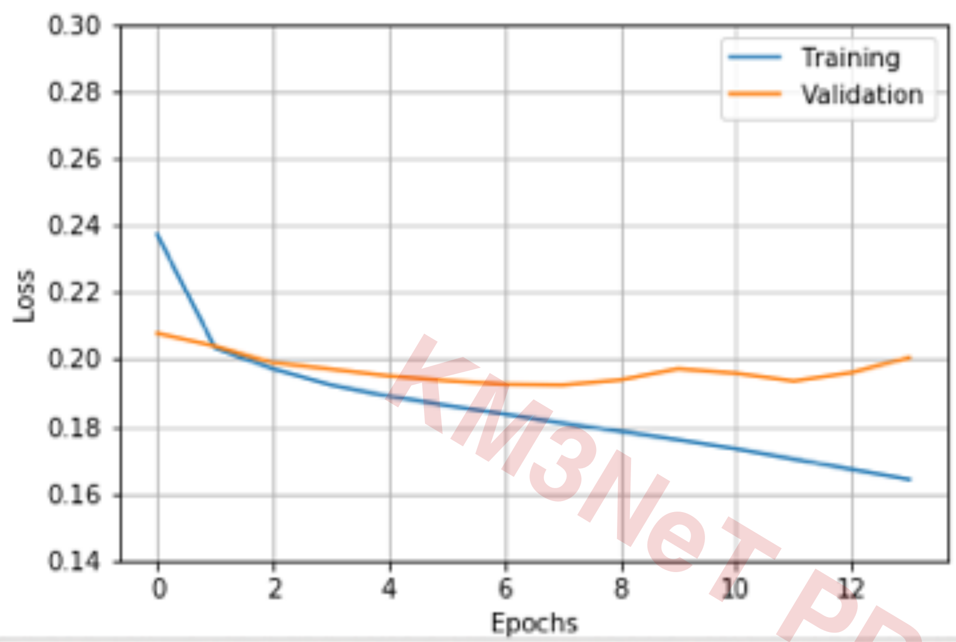
XY View

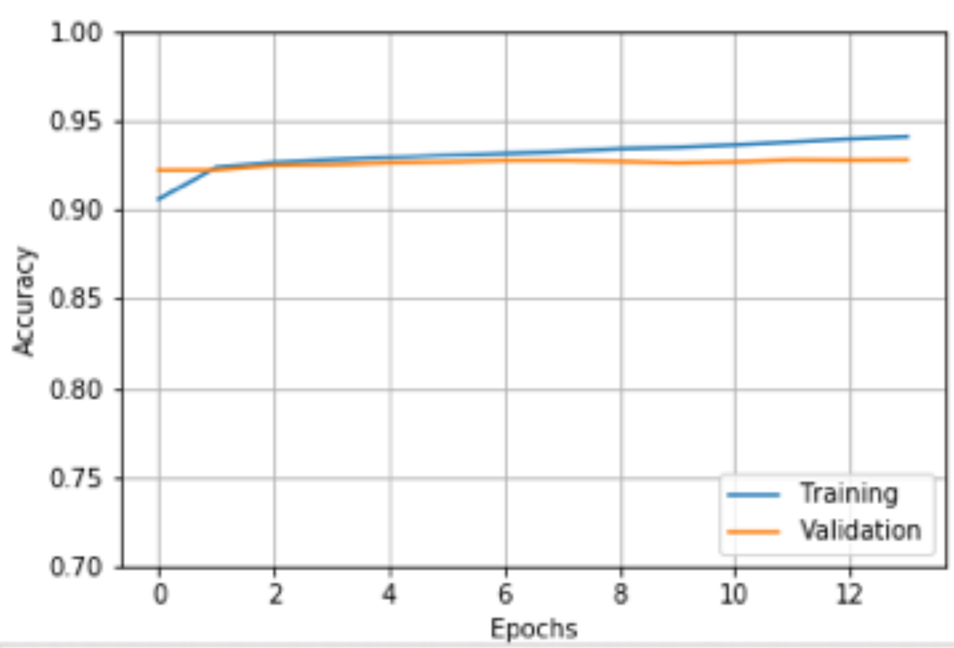


2.5 TeV

Classification results

Training history



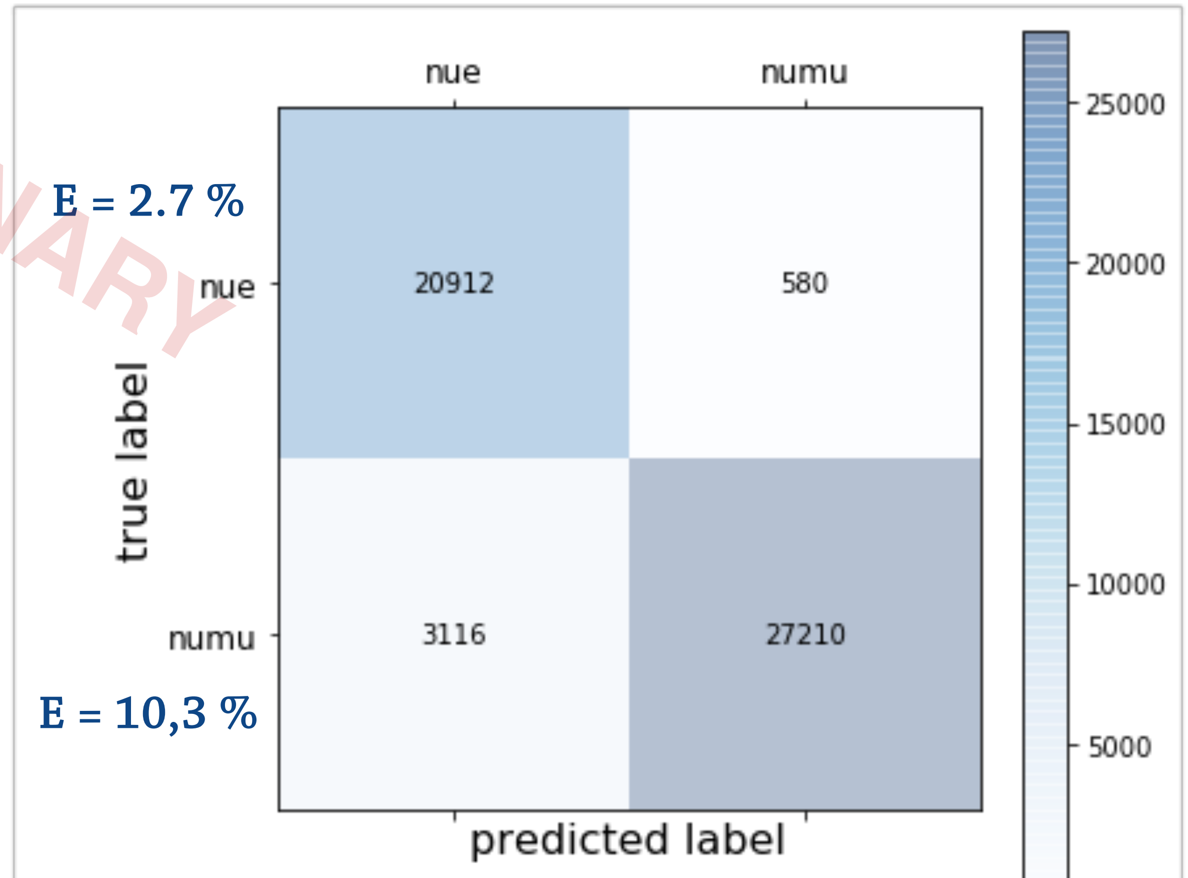


KM3NeT PRELIMINARY

Classification accuracy (on test data)

92.8 %

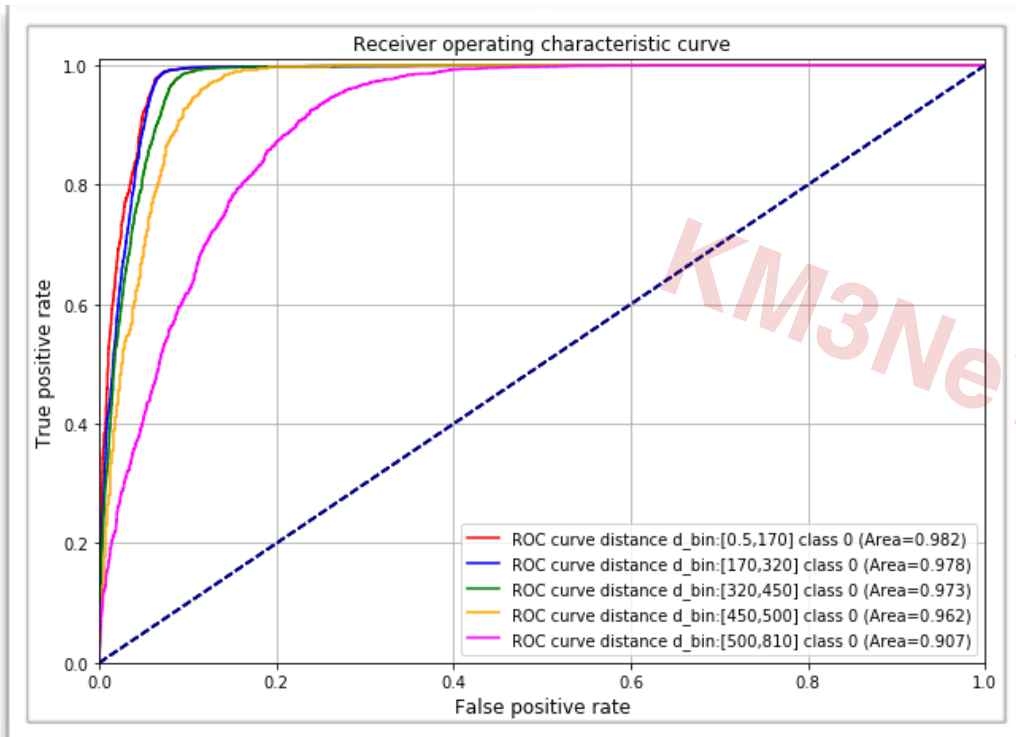
Confusion matrix (on test data)



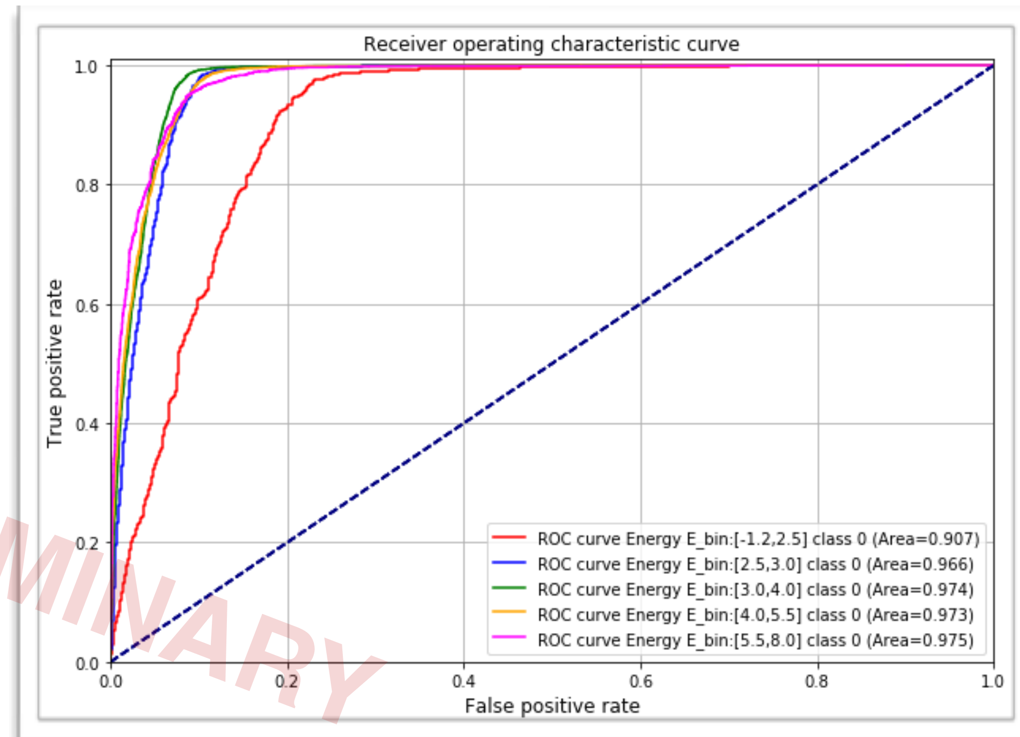
Energy and Distance dependency

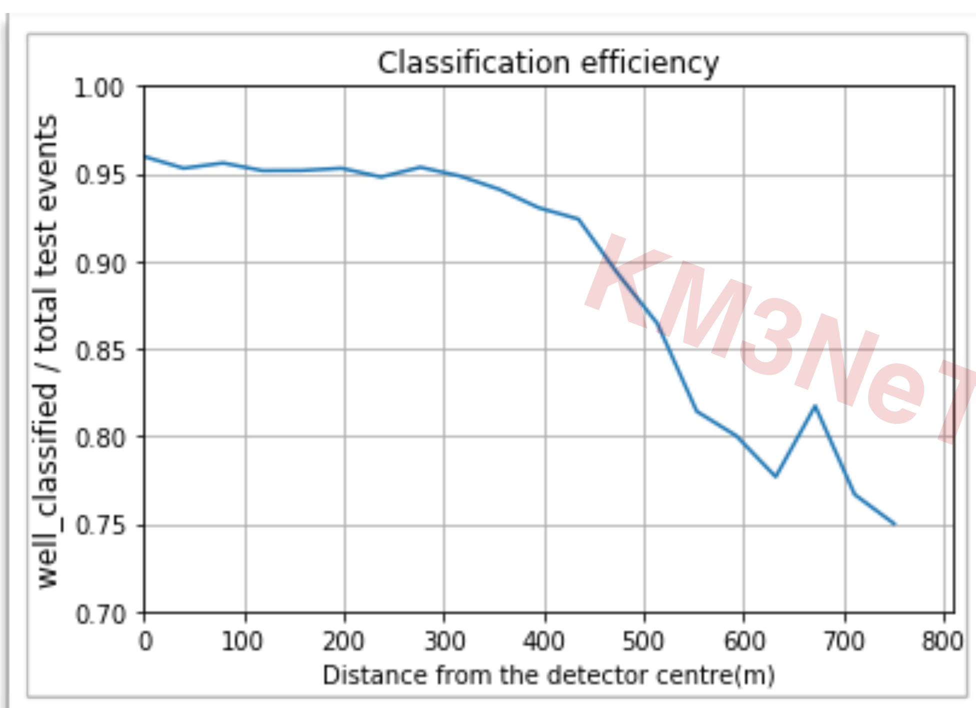
Distance bins (m)
[0.5, 170, 320, 450, 500, 810]

Energy bins (Log(E)) [GeV]
[-1.2, 2.5, 3.0, 4.0,5.5, 8.0]

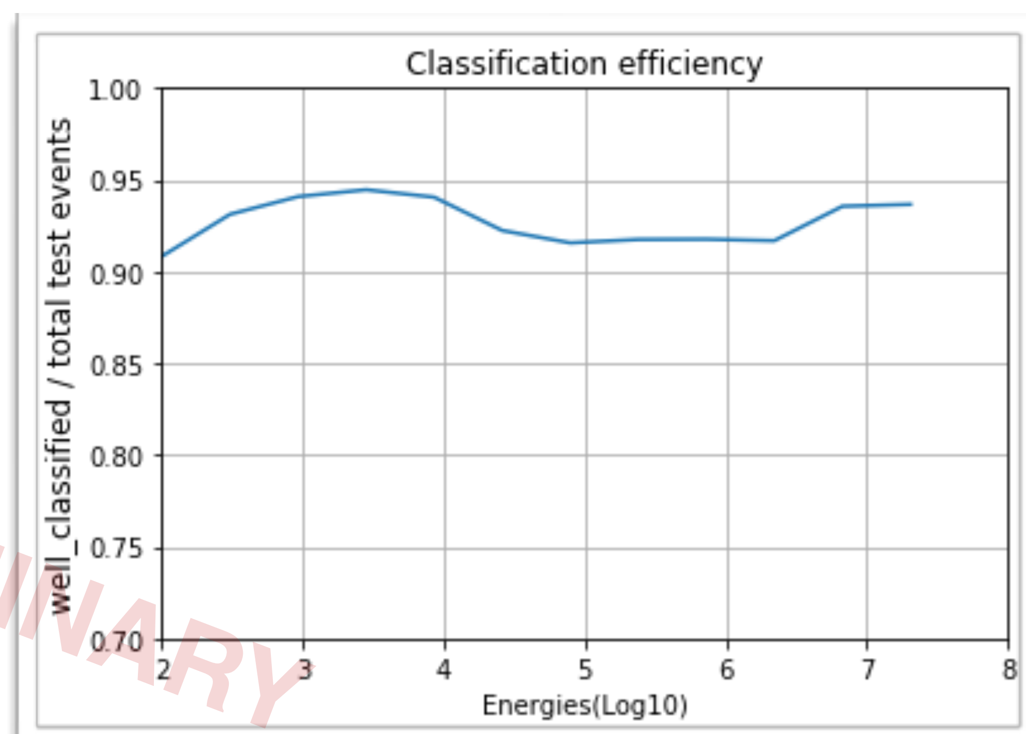


ROC curve





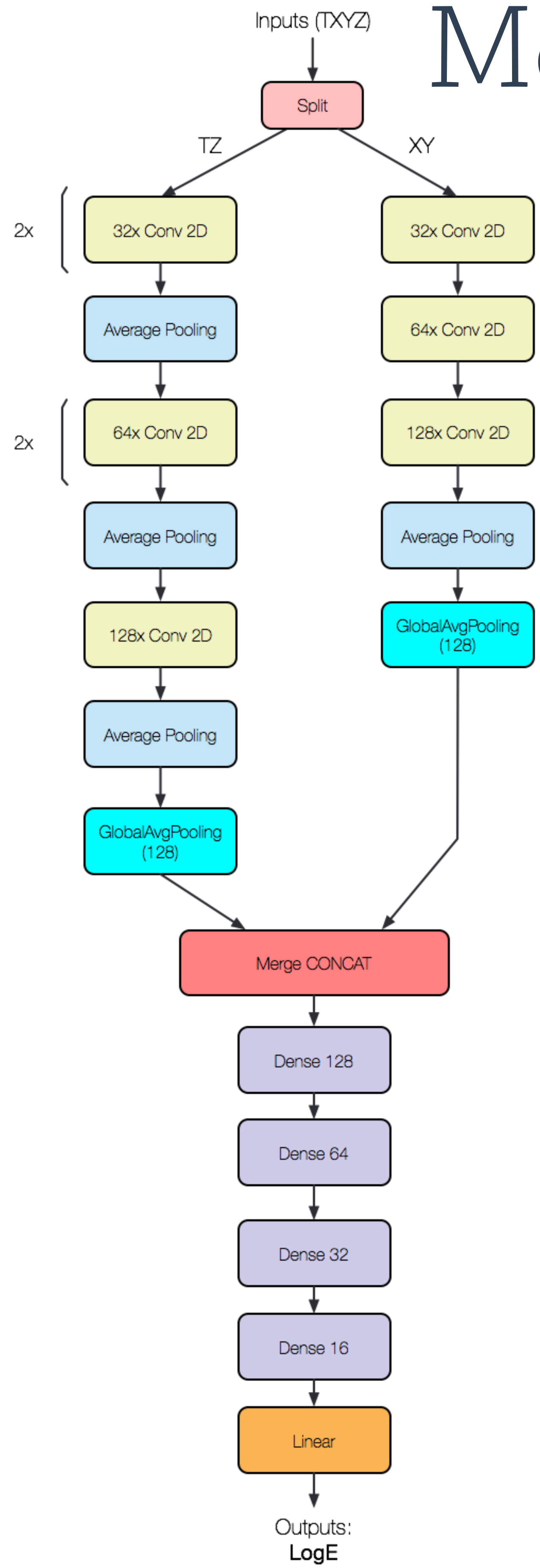
Efficiency



LEARNING TASK 3. PARTICLE ENERGY ESTIMATION

REGRESSION MODEL TO ESTIMATE NEUTRINO ENERGY (GEV)

Model architecture and Input Data



CNN model with parallel branches analysing (T,Z) and (X,Y) evolution separately, merged to extract common features, fed into multiple fully connected layers

2 Input tensors of shapes:

(T,Z): [n_samples, discrete_time_index, z_index]

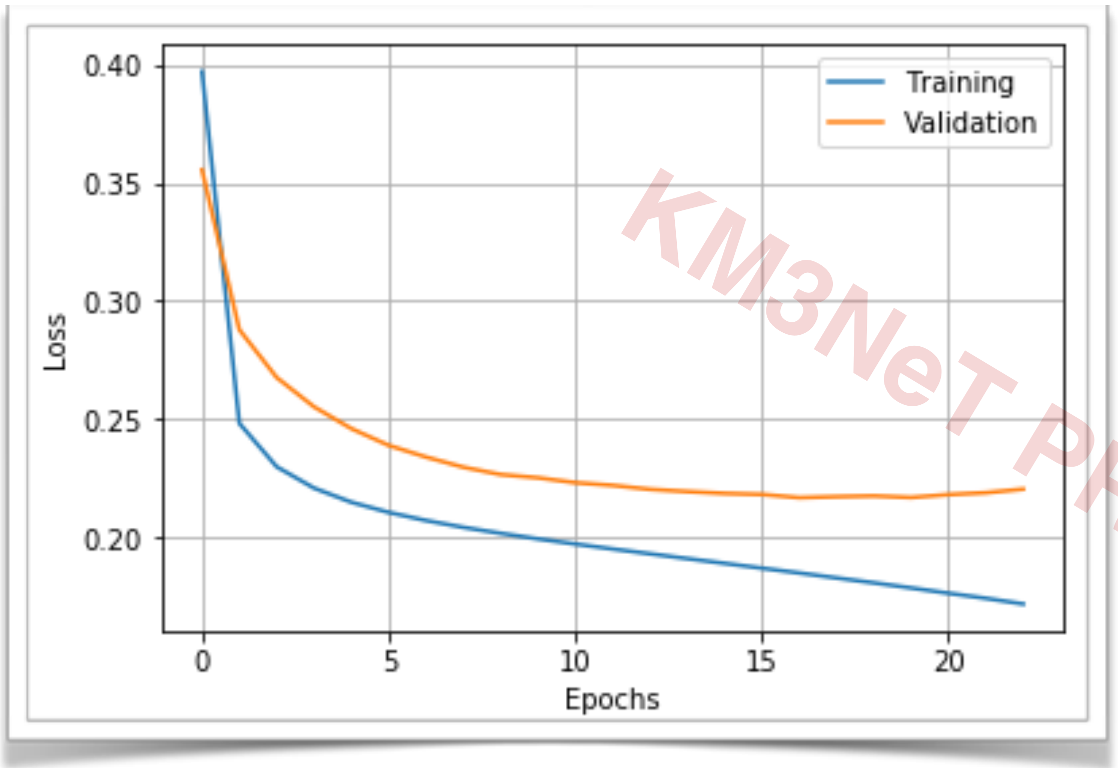
(X,Y): [n_samples, x_index, y_index]

Labels:

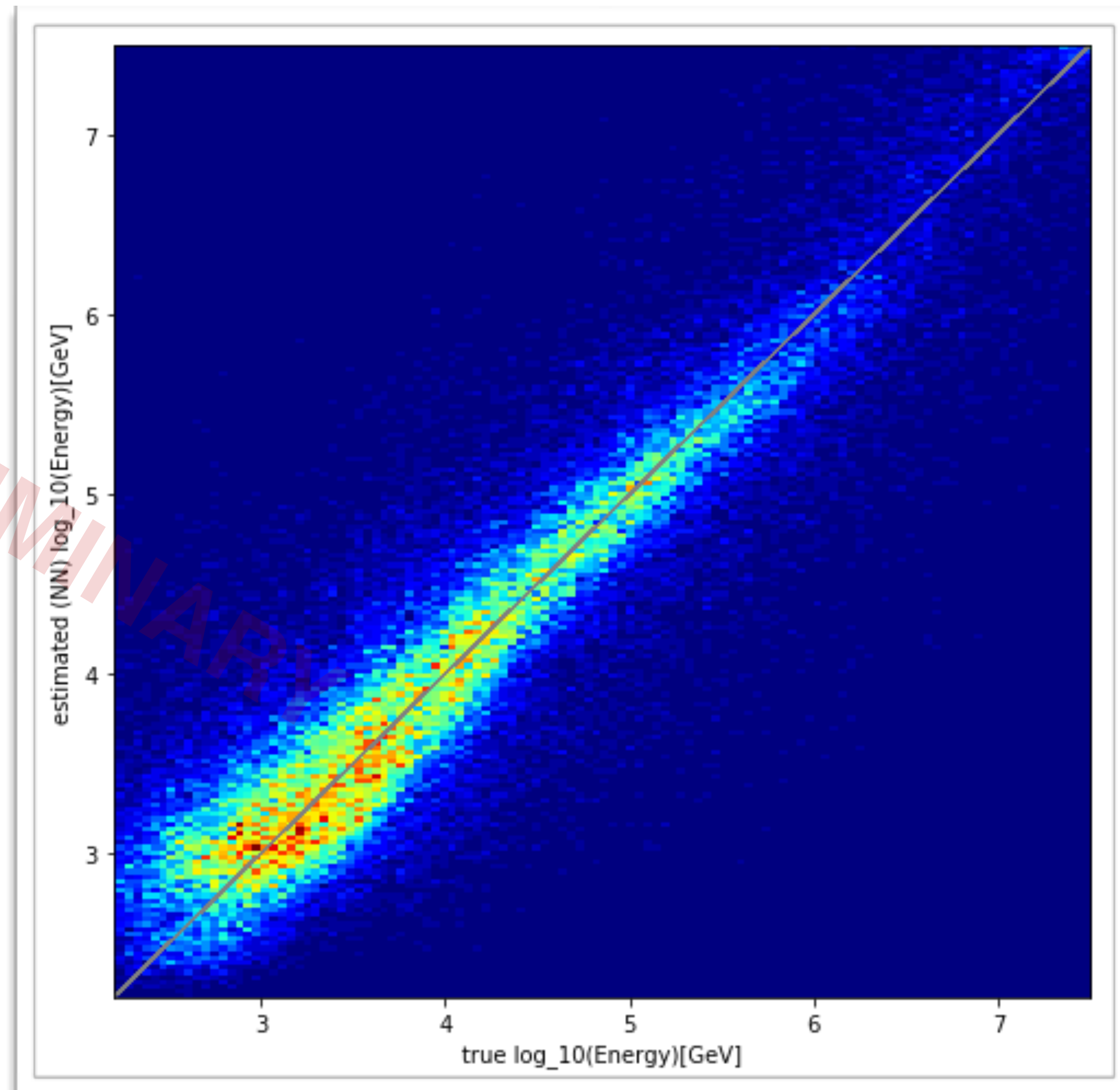
Energy: MC truth

Results

Training history



Estimated vs True Energy



test dataset ($\nu_\mu CC + \nu_e CC$ events)

Mean Squared Error (on test data)

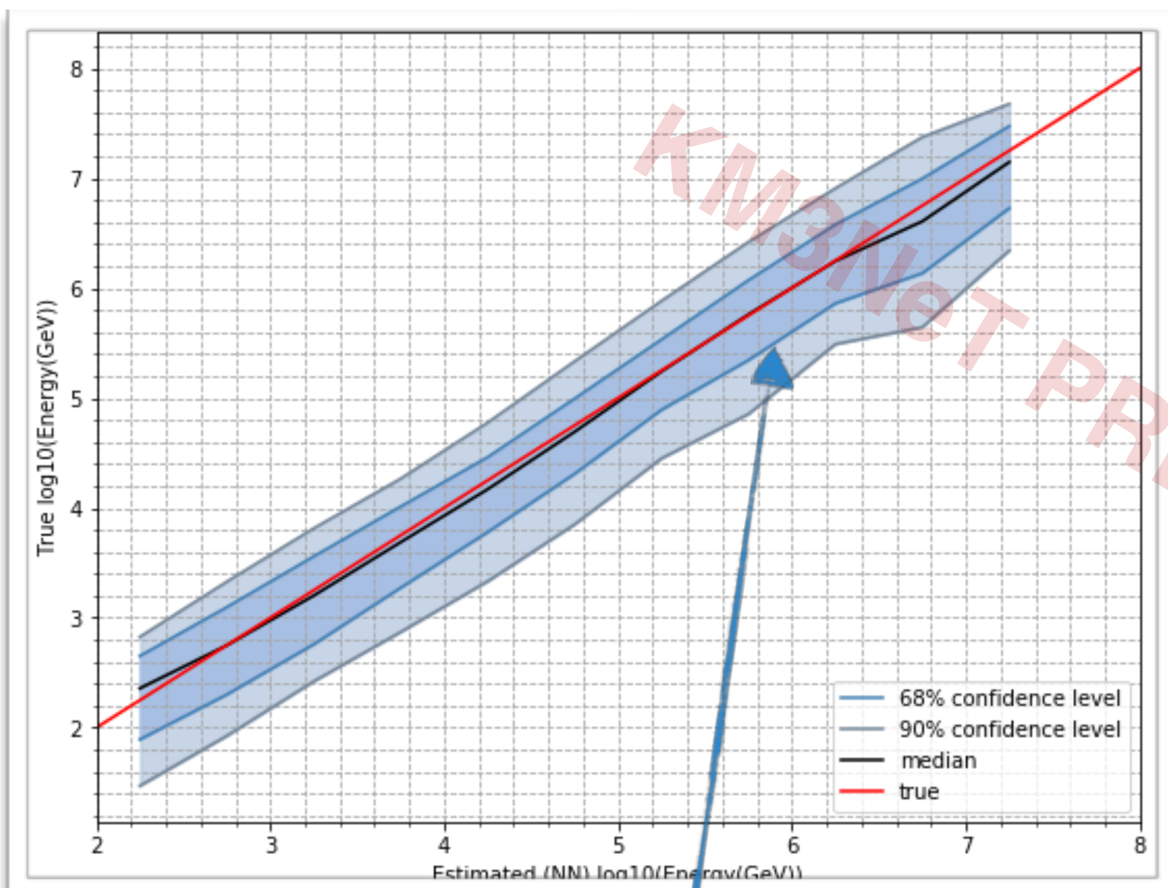
0.22

r^2 score (on test data)

0.84

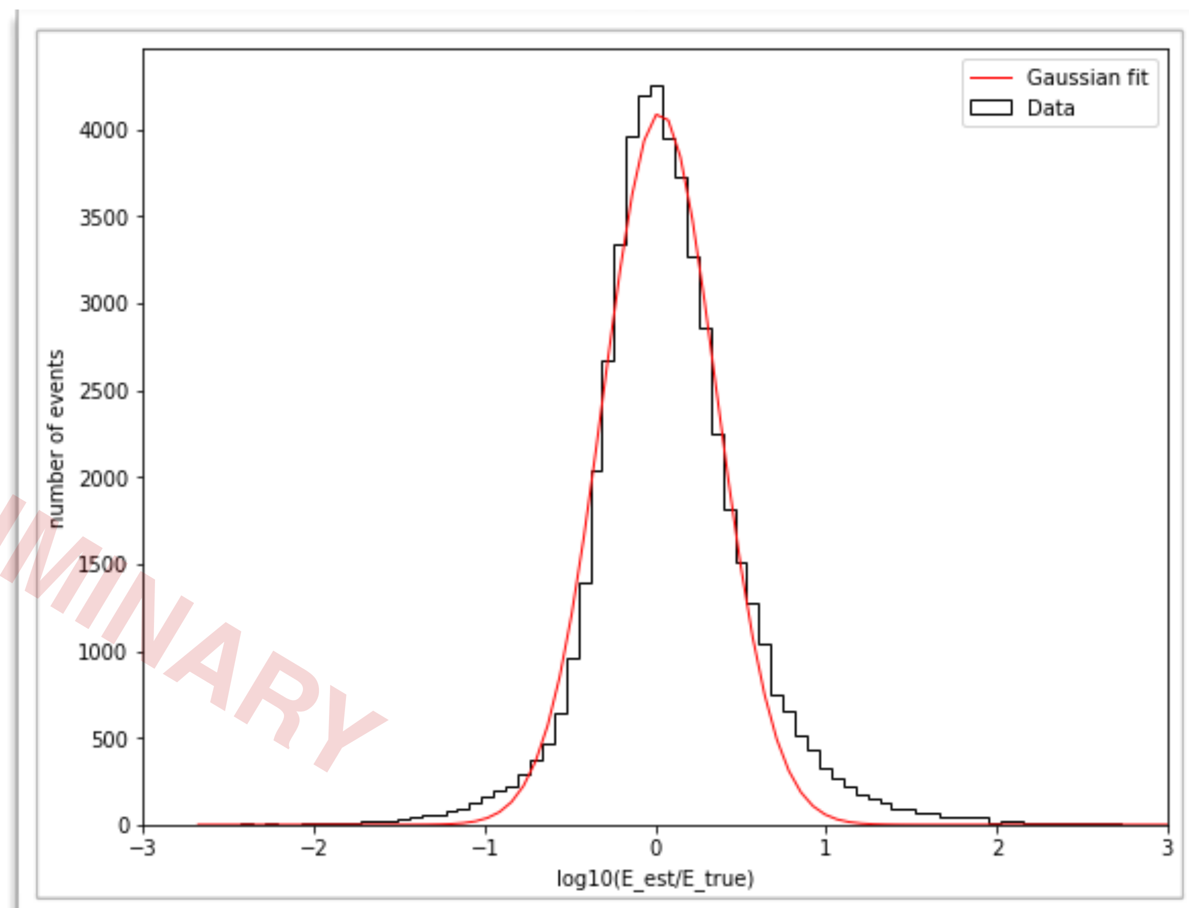
Results

True vs Neural Network Estimated Energy



lack of statistics at high energy

Energy resolution



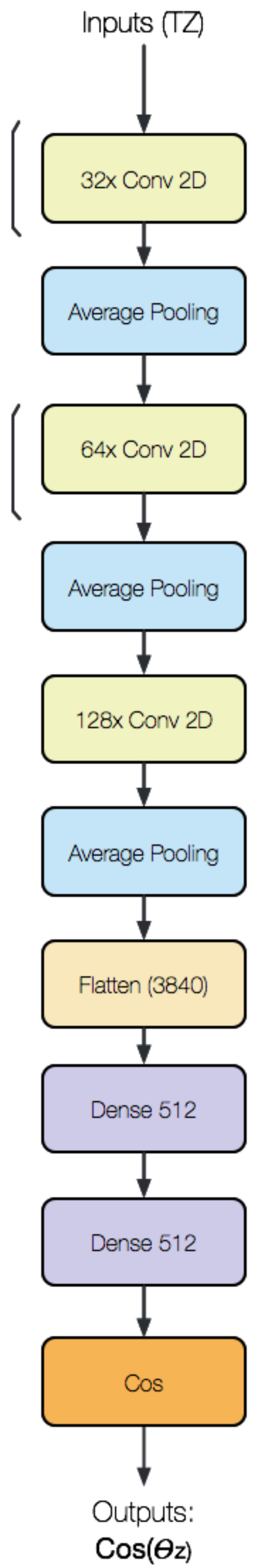
Gaussian fit w/ μ = 0.03, σ = 0.33

test dataset ($\nu_\mu CC + \nu_e CC$ events)

LEARNING TASK 4. DIRECTION ESTIMATION (Z)

REGRESSION MODEL TO ESTIMATE Z-COMPONENT OF THE NEUTRINO DIRECTION

Model architecture and Input Data



CNN model to analyse (T,Z) evolution and predict $\cos(\theta_z)$ value

Input tensor reshaped to :

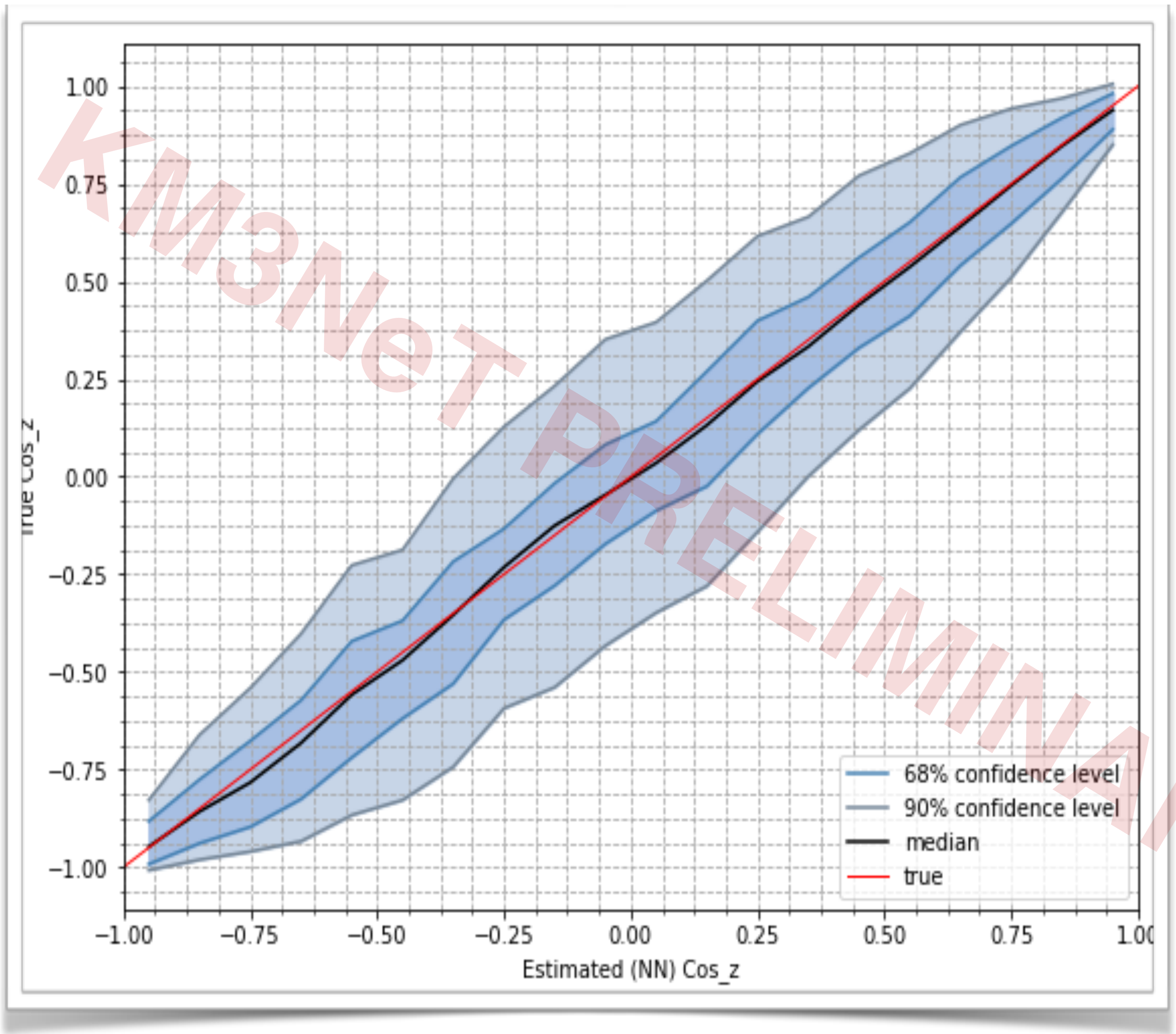
(T,Z): [n_samples, discrete_time_index, z_index]

Labels

$\cos(\theta_z)$: MC truth

Results

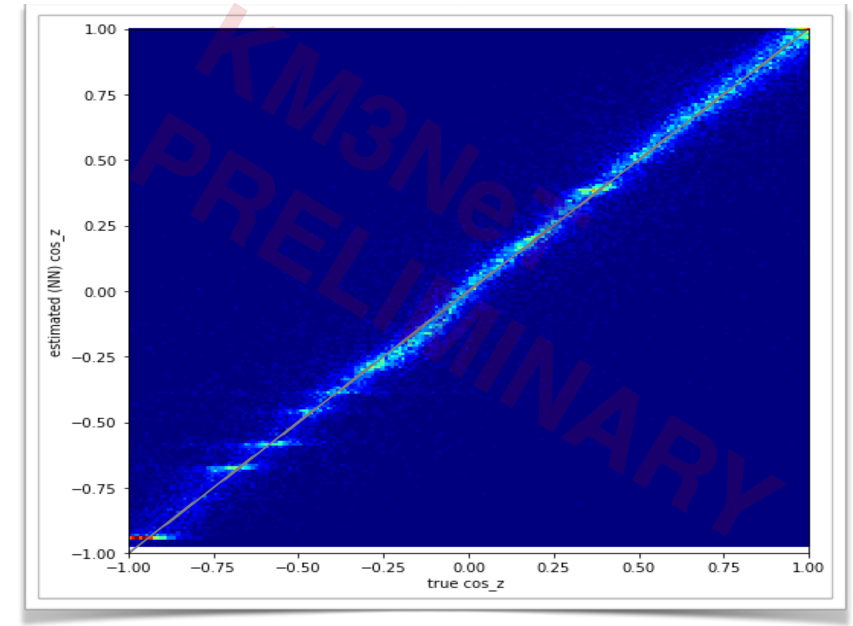
True vs NN Estimated direction (Z)



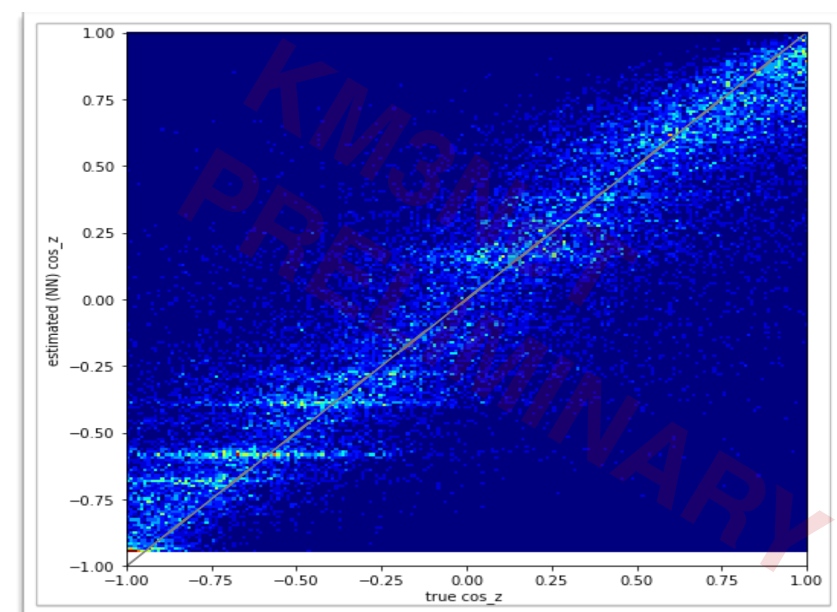
test dataset ($\nu_\mu CC + \nu_e CC$ events)

Mean Squared Error (on test data) **0.03**

$\nu_\mu CC$ vs. $\nu_e CC$



$\nu_\mu CC$



$\nu_e CC$

Muon neutrino event direction is better estimated w.r.t. electron neutrino direction

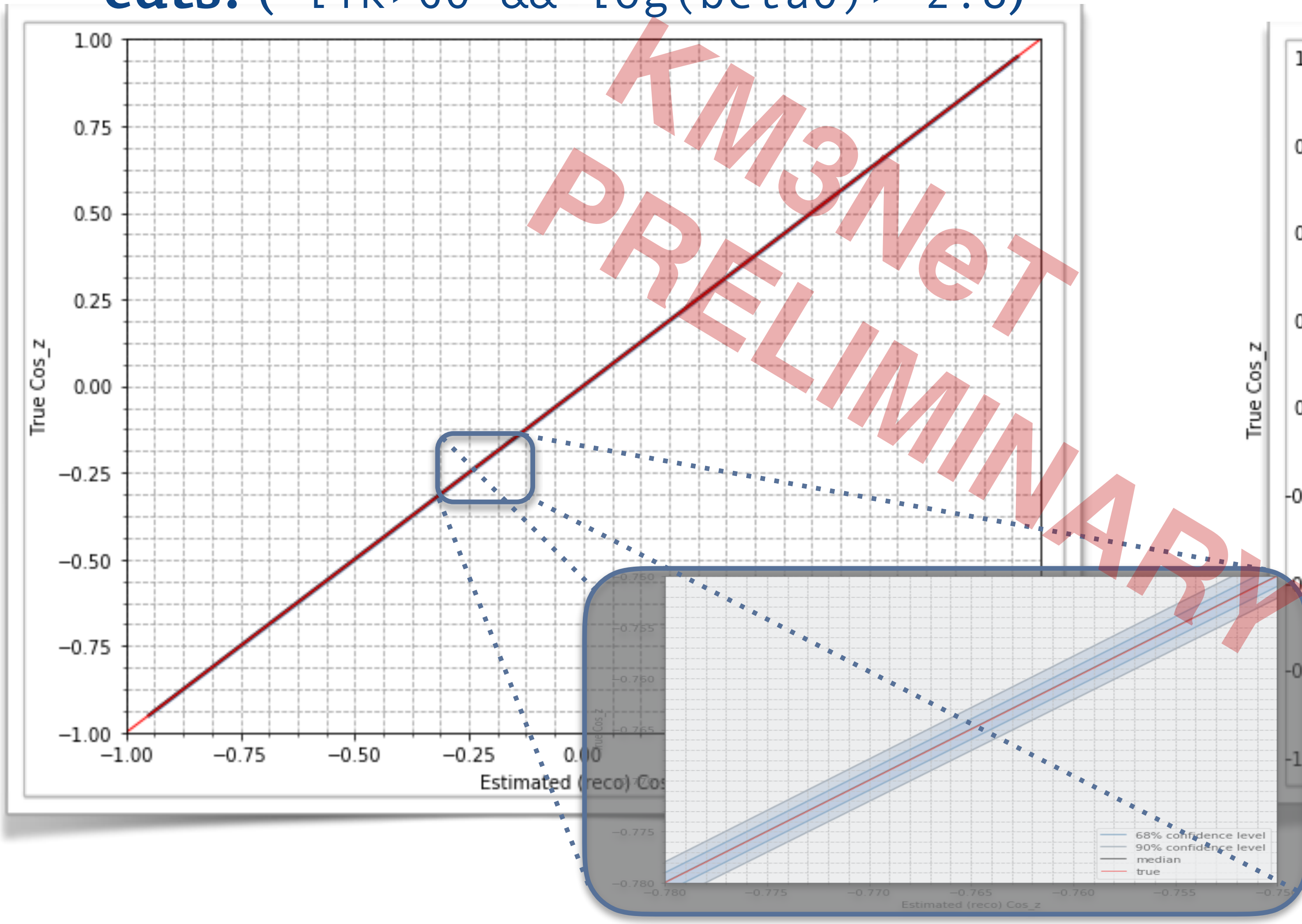
COMPARISON WITH THE OFFICIAL RECONSTRUCTION ALGORITHM

PERFORMANCES COMPARISON ON $\nu_\mu CC$ EVENTS
TO COMPARE RESULTS WITH STANDARD TRACK RECONSTRUCTION ALGORITHM, WHEN APPLICABLE

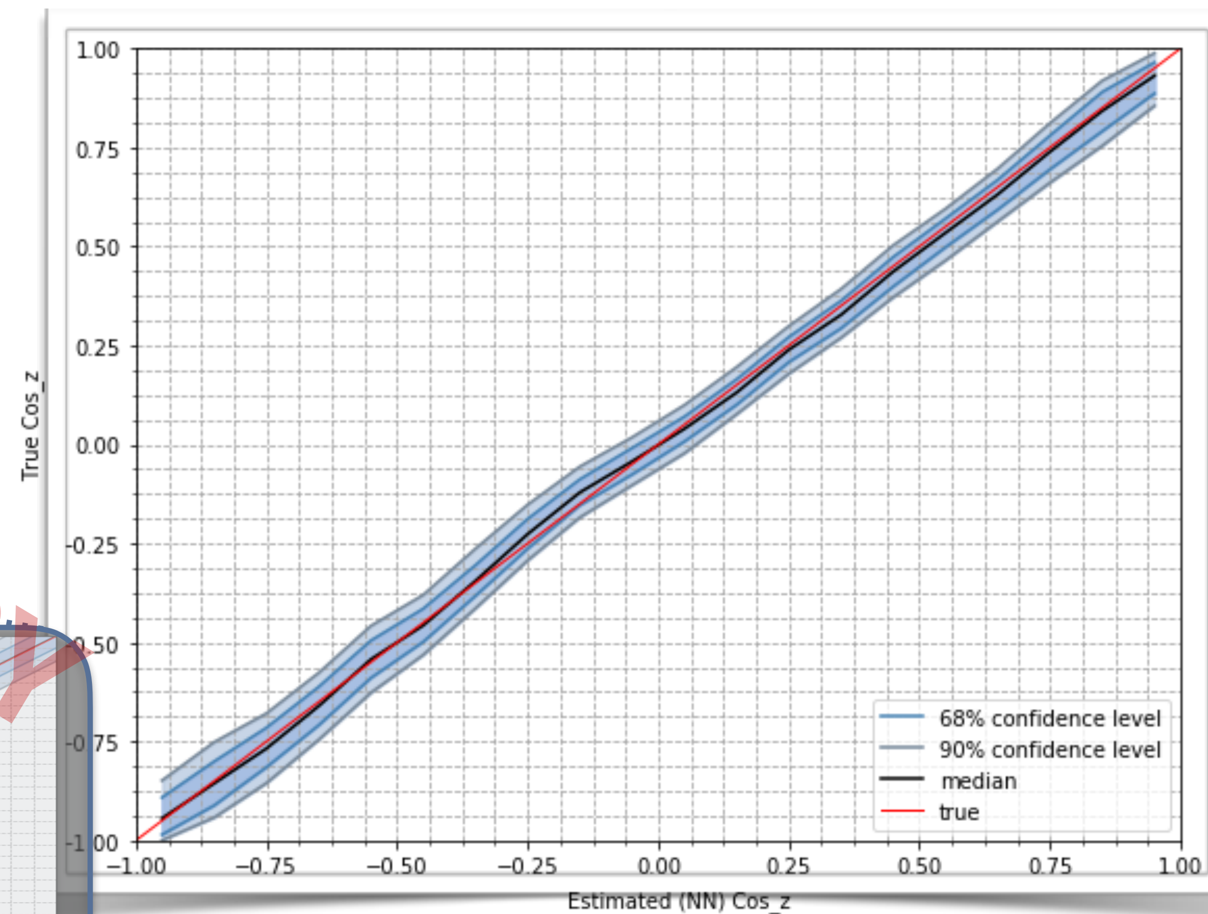
Direction estimation (Z)

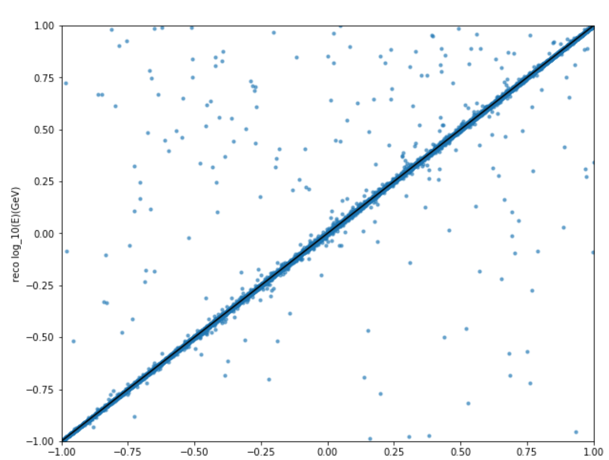
True vs JGandalf Est. direction (Z)

cuts: `(-lik>60 && log(beta0)>-2.8)`

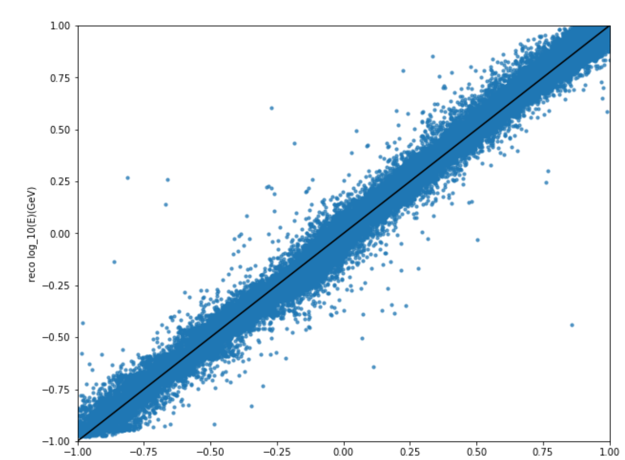


True vs CNN Est. direction (Z) on **the same** $\nu_\mu CC$ events





Estimated vs true
Neutrino direction (Z)
scatter plots



Comparison on Up-going/Down-going classification

- apply "labels" to reconstructed events:
  - $\cos(\theta_z)>0$: "up-going"
  - $\cos(\theta_z) <= 0$: "down-going"
- Compare predictions

Accuracy on up-going/down-going classification

Official reco (`JGandalf`)
(-lik>60 && log(beta0)>-2.8)

Classification Accuracy
0.998

CNN regression running test on $\nu_\mu CC$ events only

Classification Accuracy
0.987

only $\nu_\mu CC$ events (selecting well reconstructed events) with quality cuts

Deep Learning Applications for KM3NeT-ORCA

Aim:

Multipurpose classification and regression studies in ORCA

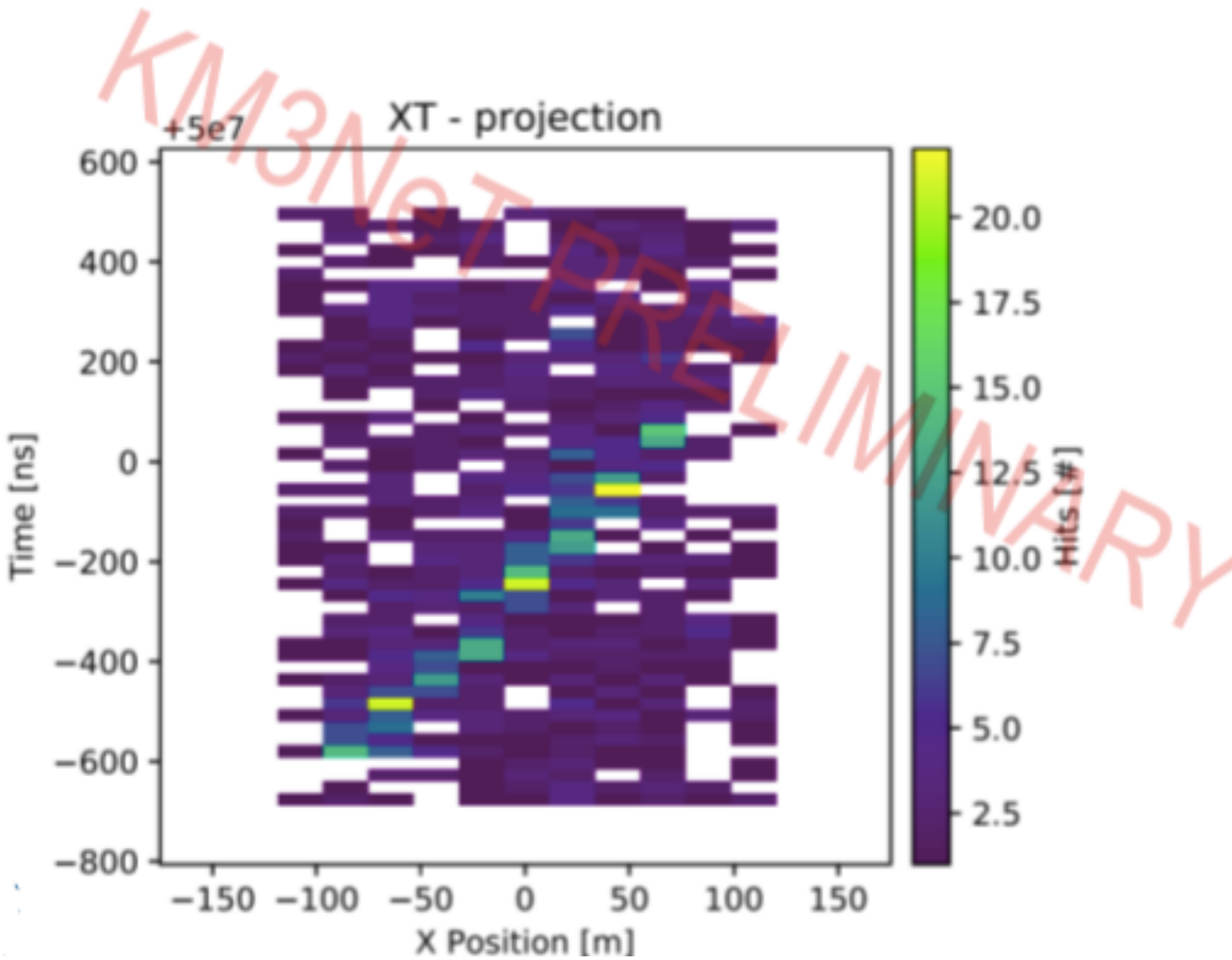
Task:

Reconstruct energy and direction, track-shower composition, PID

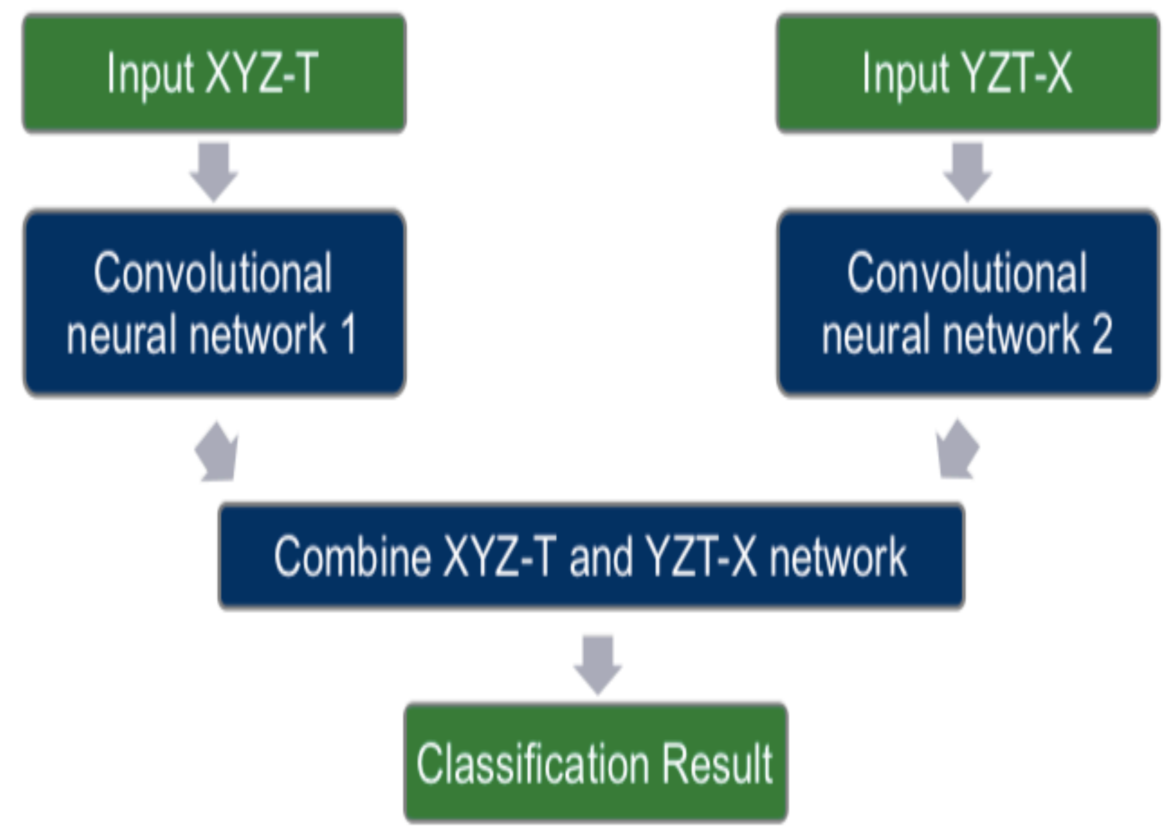
- Developed a Deep Learning track-shower classifier called OrcaNet, based on Convolutional Neural Networks

Deep Learning for ORCA with OrcaNet

- 6D information (XYZ, T, azimuth, zenith) projected to 4D subspace (neglect PMT direction)
- Time binning: 60 bins (~10 ns/bin)
- Space binning: 11x13x18
- 4 week training on HPC GPU Cluster (4xGTX 1080)



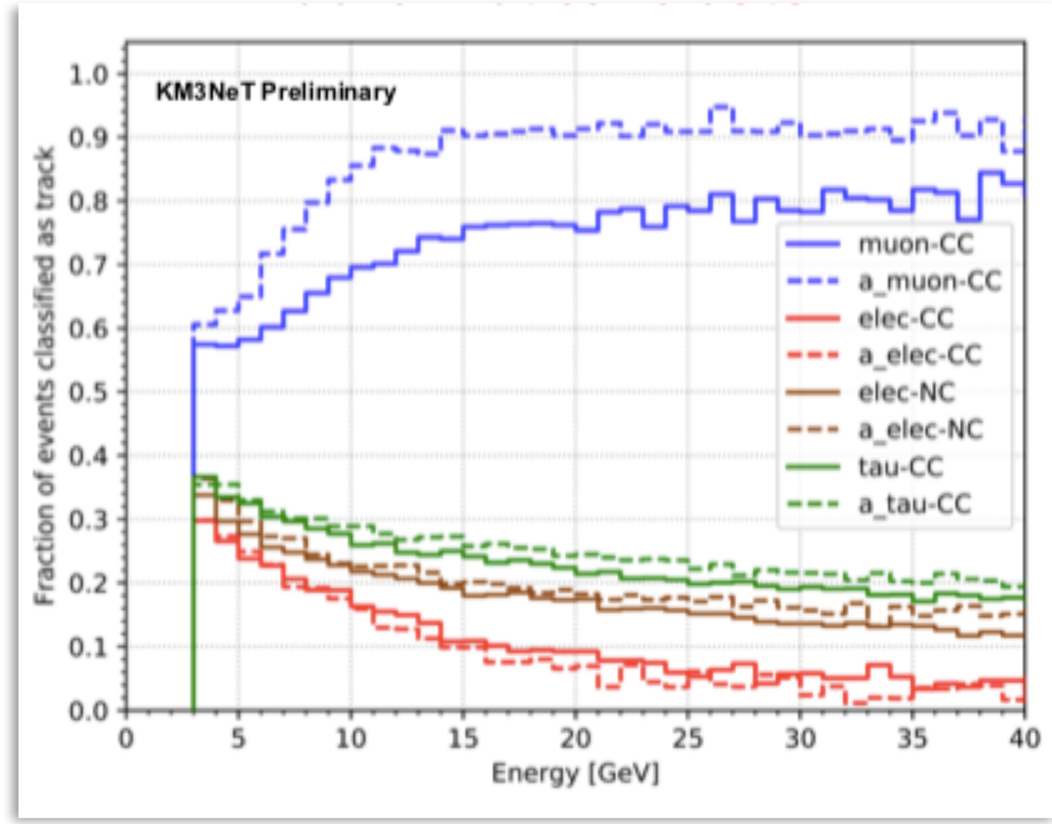
XY view of a ν_μCC event



CNN model architecture

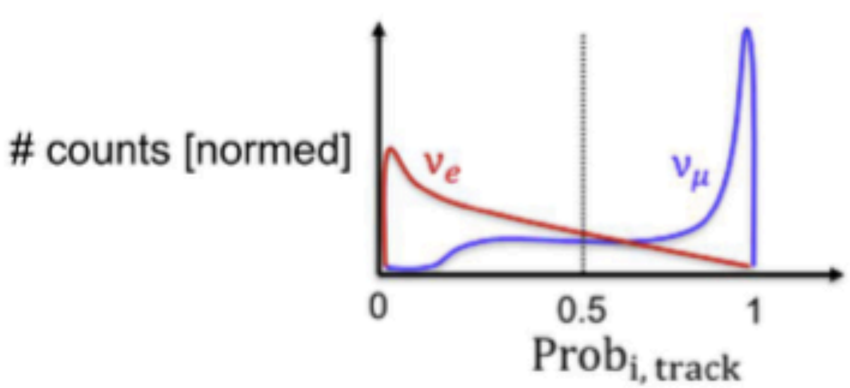
Results

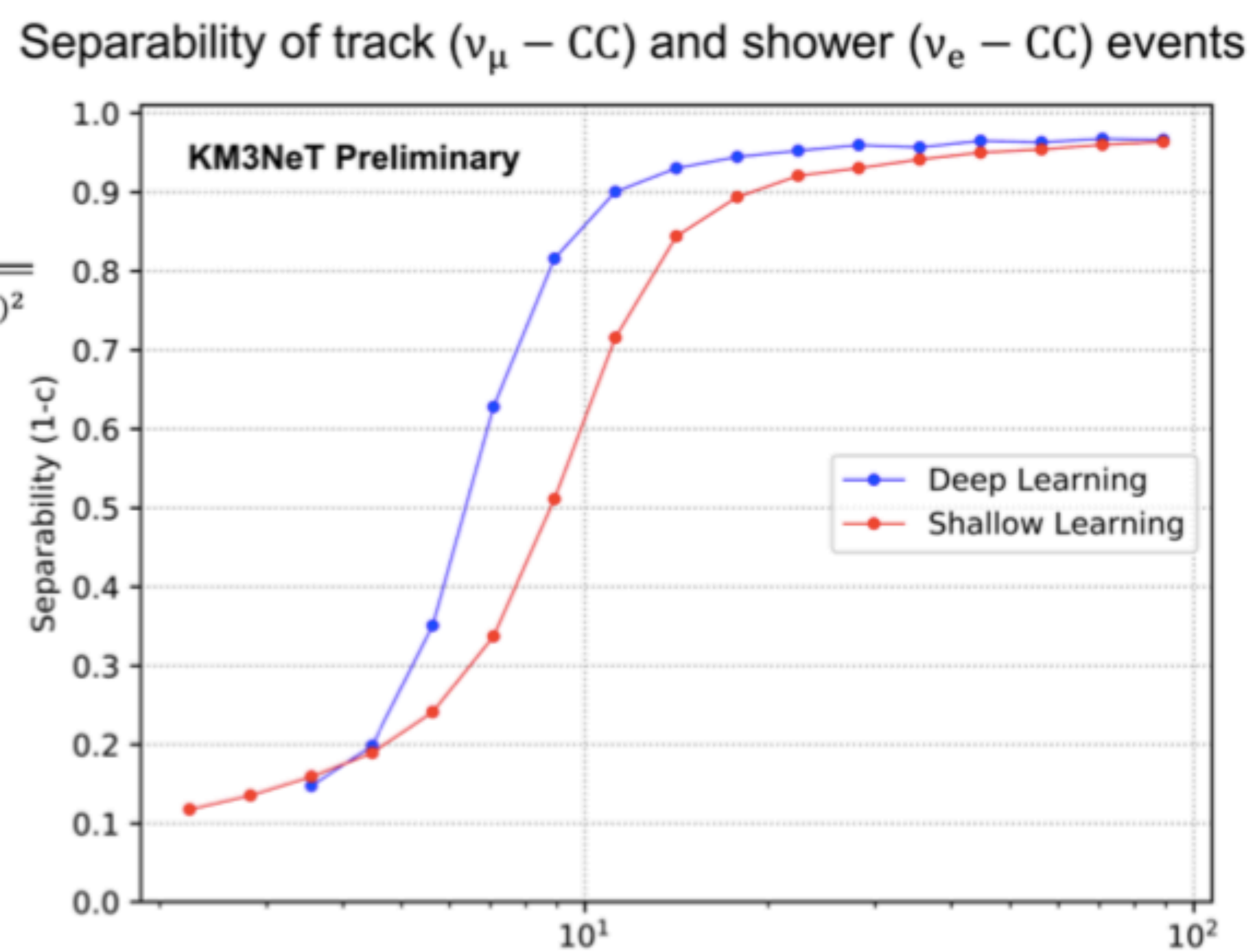
Track/shower event classification
No reconstruction output used



CNN **outperforms** RDF with human-selected features

$$S(E) = 1 - c = 1 - \frac{\sum_i P^{\nu_\mu}_{i,\,\mathrm{track}}(E)\cdot P^{\nu_e}_{i,\,\mathrm{track}}(E)}{\sqrt{\sum_i (P^{\nu_\mu}_{i,\,\mathrm{track}}(E))^2 \cdot \sum_i (P^{\nu_e}_{i,\,\mathrm{track}}(E))^2}}$$





Conclusions and outlook

- ML Results are promising in these **first iterations**:
    - ML provides **stable estimations,** comparable and in some cases better than official reco
    - ML does not depend on any reconstruction algorithm: **independent** event study directly from **raw data**
- Room for improvement and a lot of work to do (detailed detector description, complete reconstruction, complete flavour identification) – More results coming soon!
- KM3NeT expects to produce a DL toolset for CNN applications in ASTERICS (H2020) - Possibly portable to other event-based experiments
- KM3NeT INFN groups to contribute with ML algorithms in Task 3.4 of ESCAPE (proposal)

Thanks a lot for your kind attention

Chiara De Sio
cdesio@unisa.it