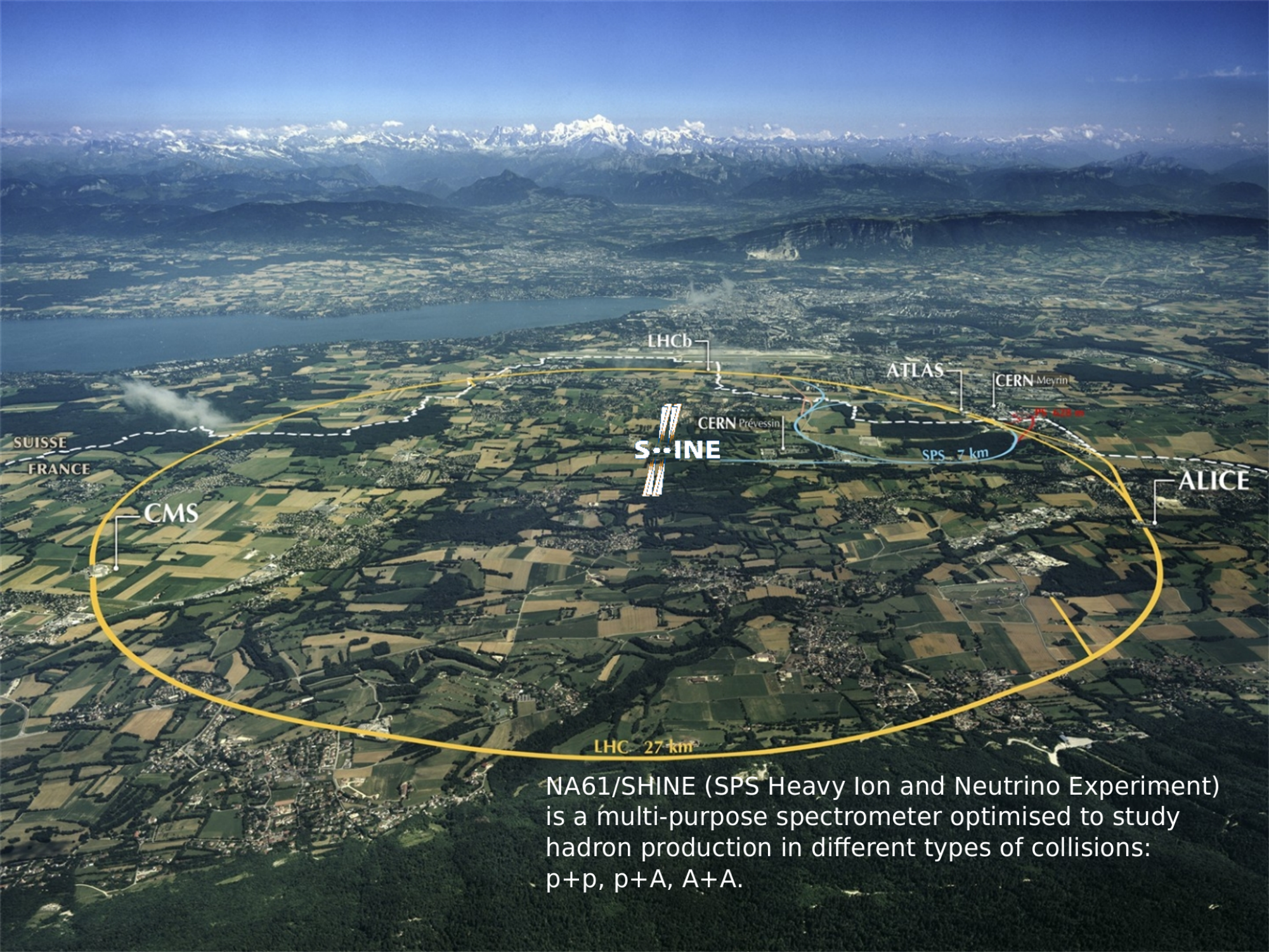
NA61/SHINE online noise filtering using machine learning methods

Anna Kawęcka

Wojciech Bryliński, Manjunath Omana Kuttan, Olena Linnyk, Janik Pawlowski,
Katarzyna Schmidt, Marcin Słodkowski, Jakub Zieliński

December 3, 2021



NA61/SHINE (SPS Heavy Ion and Neutrino Experiment) is a multi-purpose spectrometer optimised to study hadron production in different types of collisions: p+p, p+A, A+A.

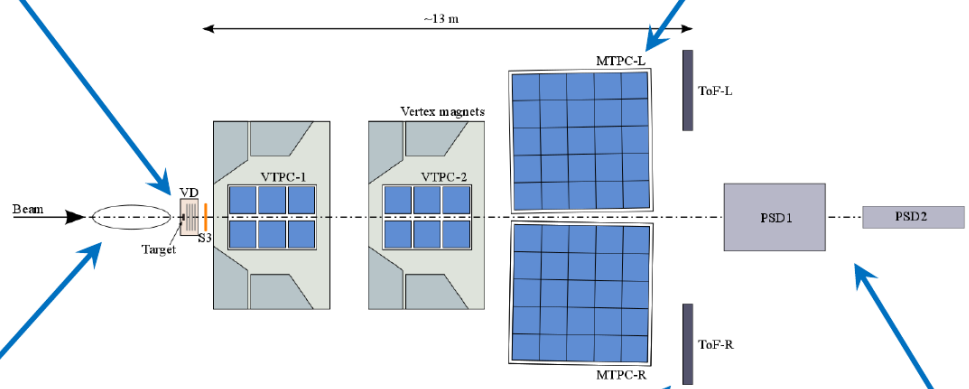
NA61/SHINE upgrade

Main goal: 10 fold increase of data taking rate up to 1 kHz

➡ **The fast and efficient tool for online noise filtering is needed**

Construction of Vertex Detector (VD) for D^0, $\overline{D^0}$ decay reconstruction

Replacement of the TPC read-out electronics to increase data rate to 1 kHz



New trigger and data acquisition system

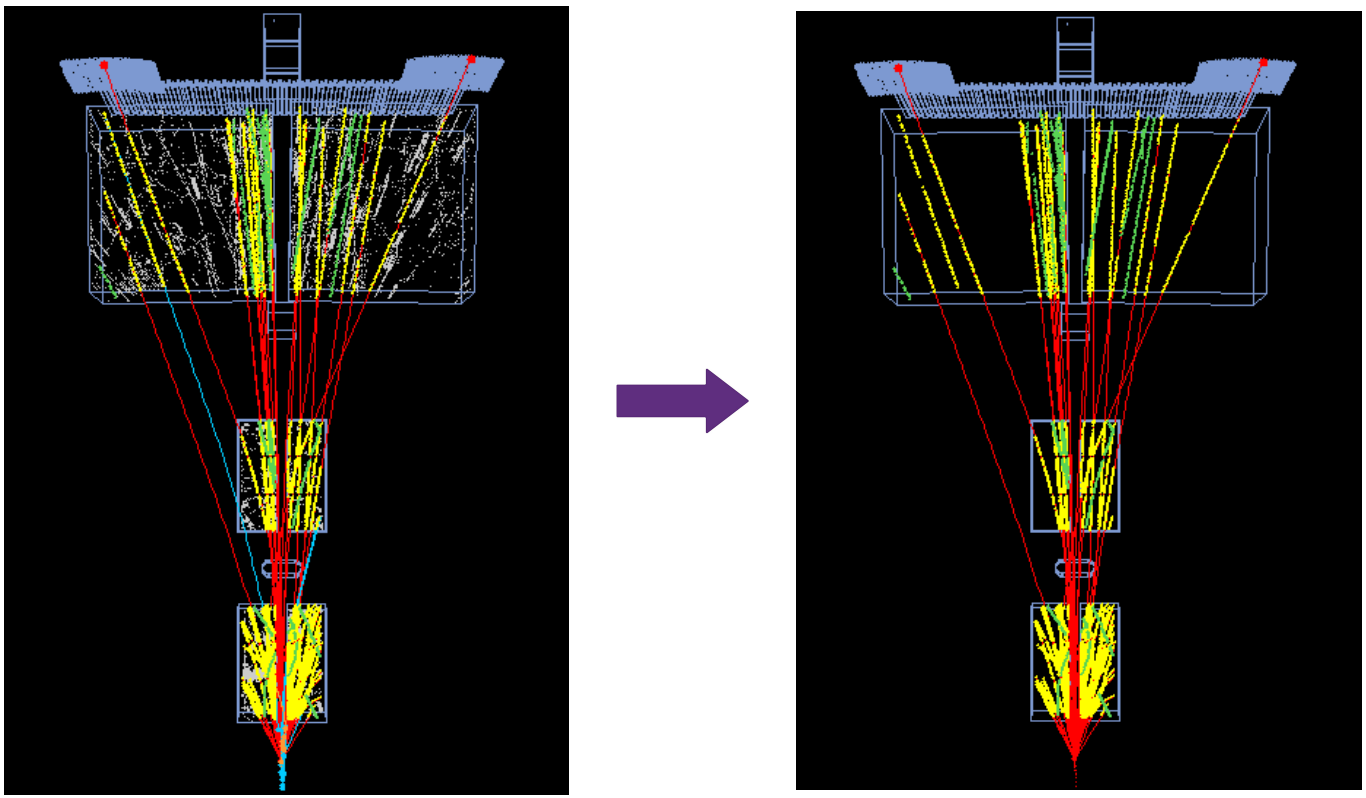
New Time-of-Flight detectors

Upgrade of Projectile Spectator Detector

NA61/SHINE TPC clusters

- TPC noisy clusters:
  - 40-50% of clusters for Pb+Pb
  - up to 70% for small systems

The standard method of track reconstruction is very slow and resource consuming.



ML approach

- **Decision tree** and **dense neural networks** (NN) based on reconstructed clusters features like total charge, positions, maxADC etc.
- **Convolutional NN** based on properties of clusters raw data.

The first step of data preparation is the reconstruction of local tracks and division of the TPC clusters into 2 groups: good clusters (belonging to track) and noisy clusters. The divided clusters are used to train the classification algorithms.

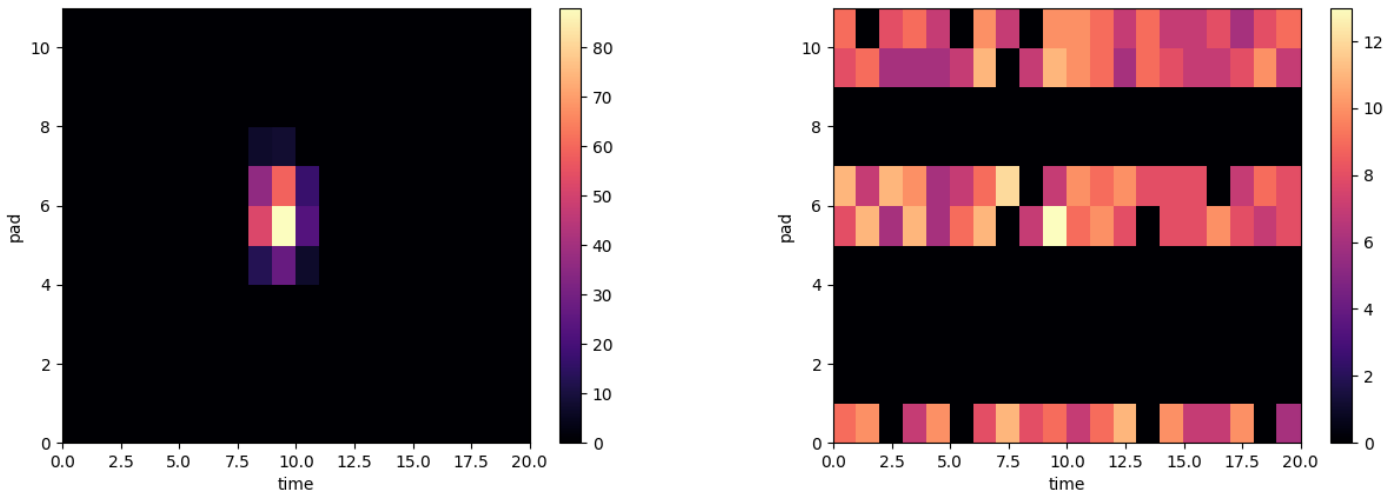


Figure: Example of good (left) and noisy cluster (right).

Confusion matrix

The main goal is to keep the **high accuracy** of good clusters recognition and keep the fraction of good clusters classified as noise (*false negatives*) as low as possible.

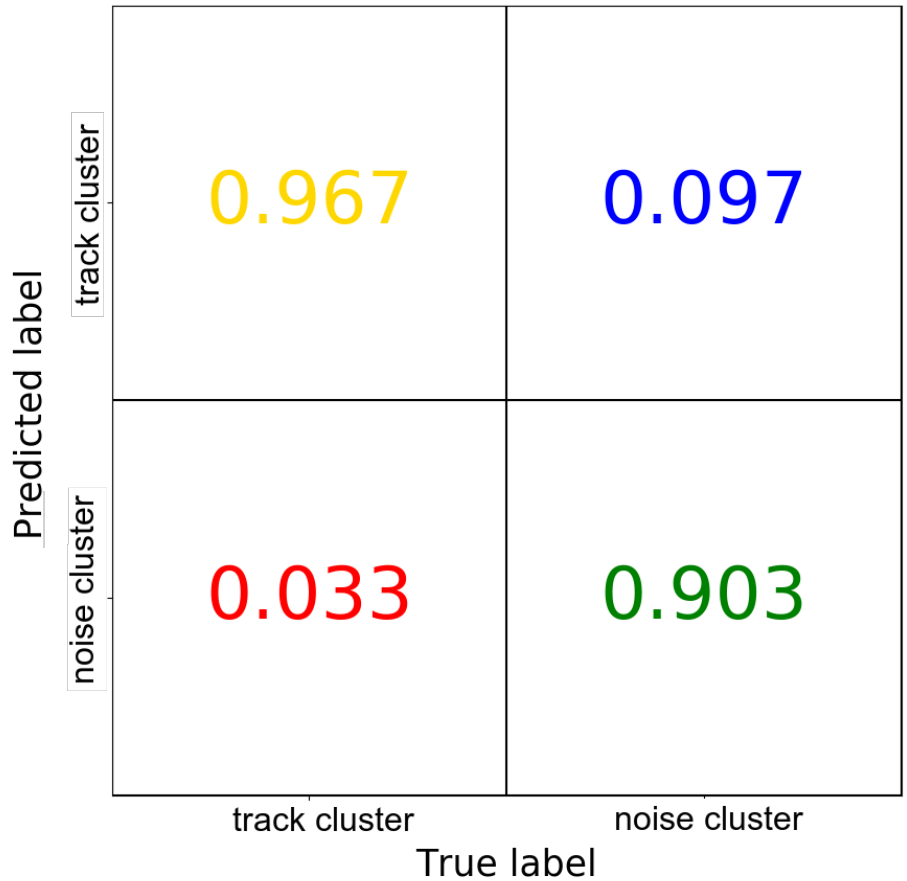


Figure: Example confusion matrix for convolutional NN.

Prediction results

The result of the classification for both methods is the probability that a given cluster was reconstructed from noise. If the probability is higher than a threshold, the cluster is removed from the data. The threshold can be selected in a way to minimise the loss of track clusters and maximise the reduction of noise.

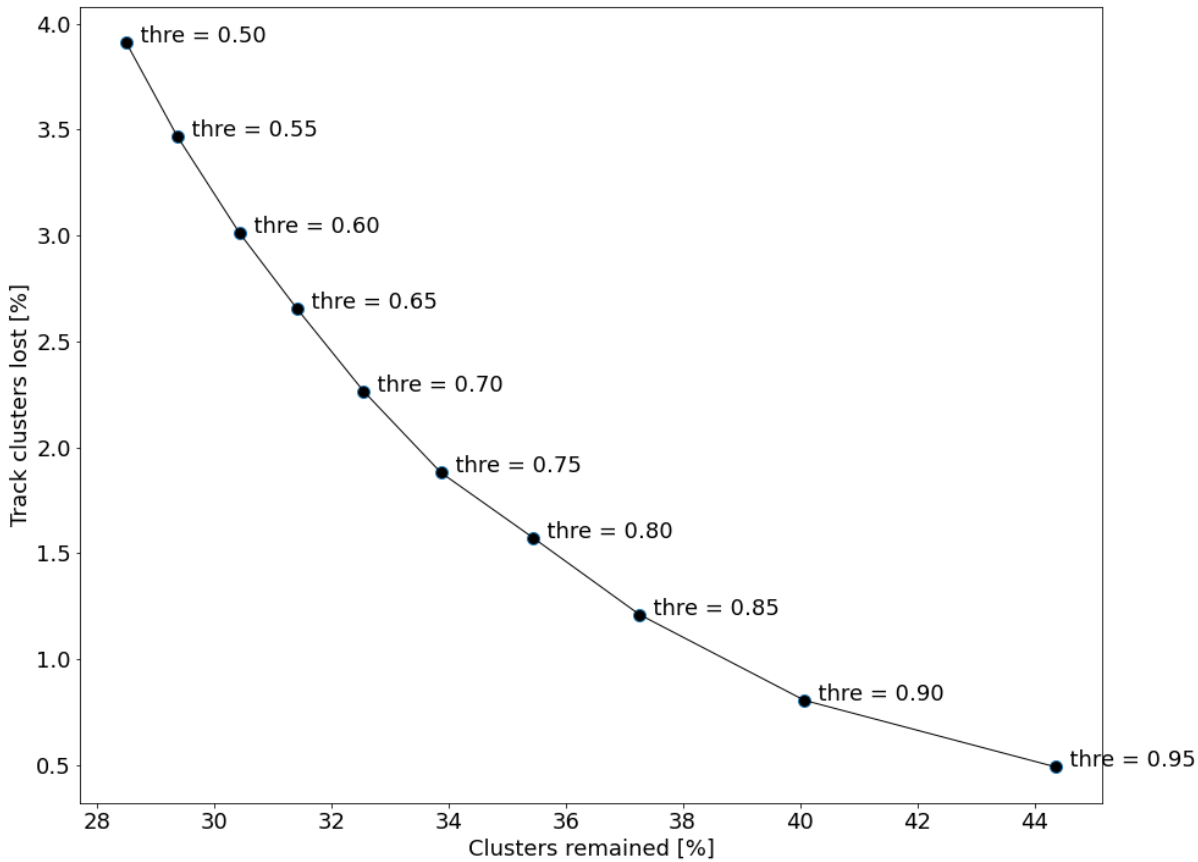


Figure: Lost track clusters as a function of remained clusters for decision tree.

Confusion matrix visualization

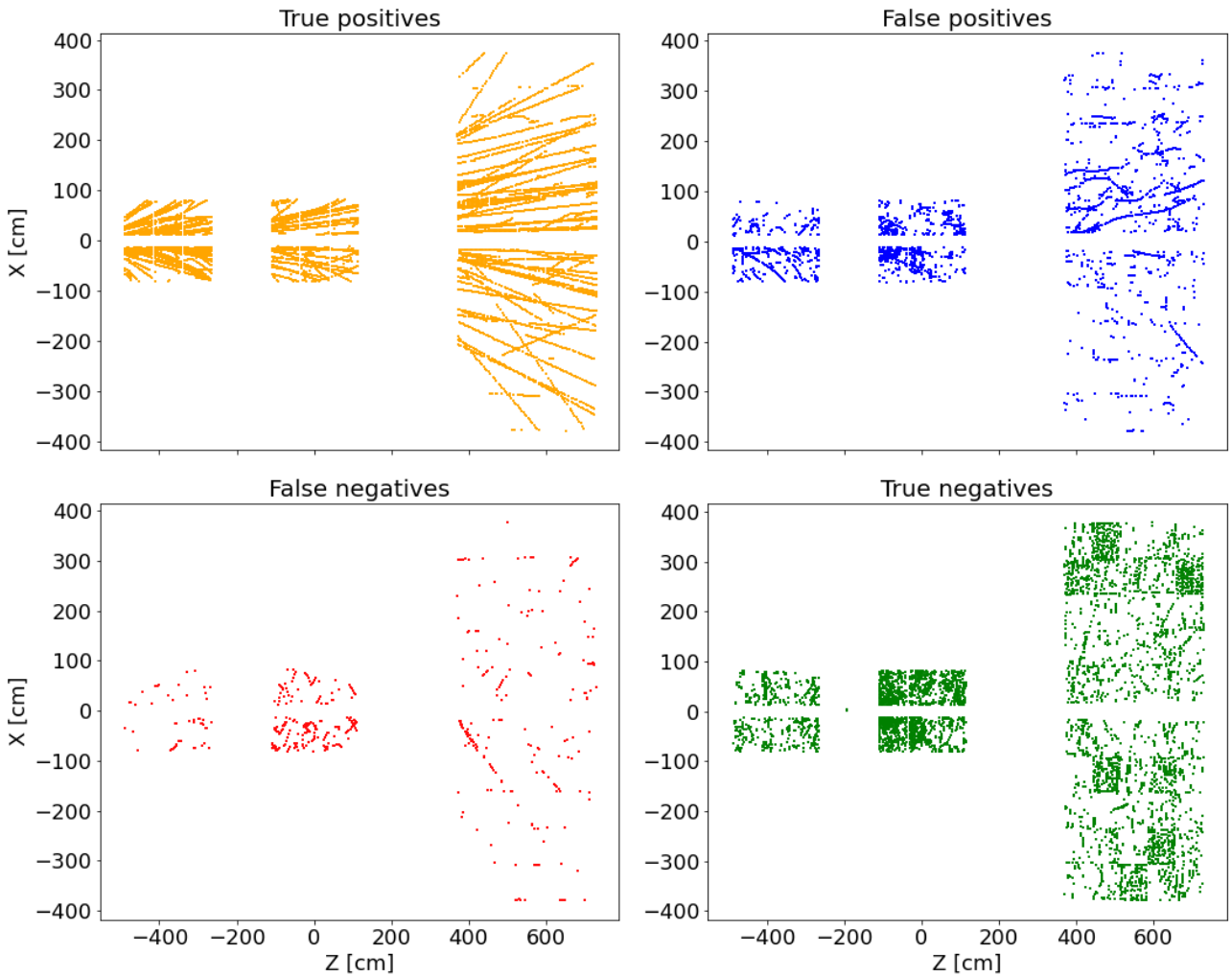


Figure: Example confusion matrix for decision tree.

Conclusions

- NA61/SHINE will need online noise filtering tool for future data taking and machine learning approach is a very efficient and robust solution.
- Two approaches were tested – using reconstructed and raw clusters information and both give good results with **prediction accuracy over 90%**.
- The algorithms are being implemented in the data acquisition (DAQ) software as online tools and the most efficient will be used during data taking.

All details of the implemented algorithms together with the comparison of results can be found here:

Thank you and have a shiny day!