Coincidence classification in the large field-of view J-PET scanners with machine learning methods

Konrad Klimaszewski, Wojciech Krzemień,
Jakub Baran, Lech Raczyński, Aldona Spirzewska

WMLQ2022 - 14.09.2022

European Funds Smart Growth
Republic of Poland
European Union European Regional Development Fund

Outline

- EuroHPC PL project
- PET imaging and Jagiellonian-PET
- Classification of PET events using ML

European Funds Smart Growth
Republic of Poland
European Union European Regional Development Fund

The European High Performance Computing Joint Undertaking (EuroHPC JU)

EuroHPC
Joint Undertaking

https://eurohpc-ju.europa.eu/index_en

joint initiative between the EU, European countries and private partner to develop a World Class Supercomputing Ecosystem in Europe

- Supercomputers
- GPU (Graphical Processor Units) computing
- Quantum computing
- Neuromorphic computing

The European High Performance Computing Joint Undertaking (EuroHPC JU)

EuroHPC Joint Undertaking

https://eurohpc-ju.europa.eu/index_en

joint initiative between the EU, European countries and private partner to develop a World Class Supercomputing Ecosystem in Europe

- Supercomputers
- GPU (Graphical Processor Units) computing
- Quantum computing
- Neuromorphic computing

EuroHPC PL

https://www.eurohpc.pl/

EuroHPC PL – development of the specialized infrastructure for the exascale computations addressing the key challenges for the Polish society, scientific community and the economy.

Consortium of 7 Polish institutions:
ACK Cyfronet AGH, PCSS, CI TASK, WCSS, NCBJ, IITiS PAN and CFT PAN.

European Funds Smart Growth
Republic of Poland
European Union European Regional Development Fund

The European High Performance Computing Joint Undertaking (EuroHPC JU)

EuroHPC Joint Undertaking

https://eurohpc-ju.europa.eu/index_en

joint initiative between the EU, European countries and private partner to develop a World Class Supercomputing Ecosystem in Europe

- Supercomputers
- GPU (Graphical Processor Units) computing
- Quantum computing
- Neuromorphic computing

EuroHPC PL

https://www.eurohpc.pl/

EuroHPC PL – development of the specialized infrastructure for the exascale computations addressing the key challenges for the Polish society, scientific community and the economy.

Consortium of 7 Polish institutions:
ACK Cyfronet AGH, PCSS, CI TASK, WCSS, NCBJ, IITiS PAN and CFT PAN.

Świerk Computing Centre

NATIONAL CENTRE FOR NUCLEAR RESEARCH ŚWIERK

Software platform for quantum simulations and medical imaging

https://www.ncbj.gov.pl/en/aktualne/eurohpc-pl-national-supercomputing-infrastructure-eurohpc

European Funds Smart Growth
Republic of Poland
European Union European Regional Development Fund

Quantum simulations and medical imaging software platform

Group:
- Wojciech Krzemień
- Konrad Klimaszewski
- Mateusz Bała
- Oleksander Fedoruk
- Lech Raczyński
- Tobiasz Jarosiewicz

Common API

Services

Simulators

PET Image Reconstructor

Phantom generator

Quatum emulators/ Quantum computer ⟷ **Quantum simulations**

Standard simulations

Image reco.

Quantum Imaging

GAN networks

Libraries

Qiskit

GEANT4 A SIMULATION TOOLKIT

GATE

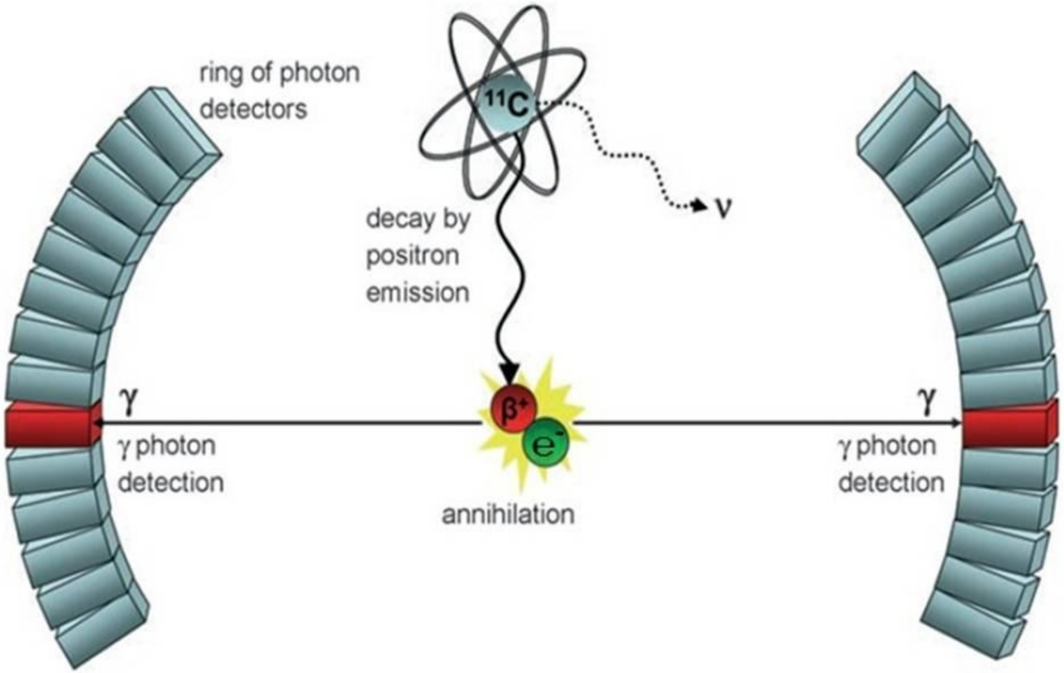
TensorFlow

TensorFlow Quantum

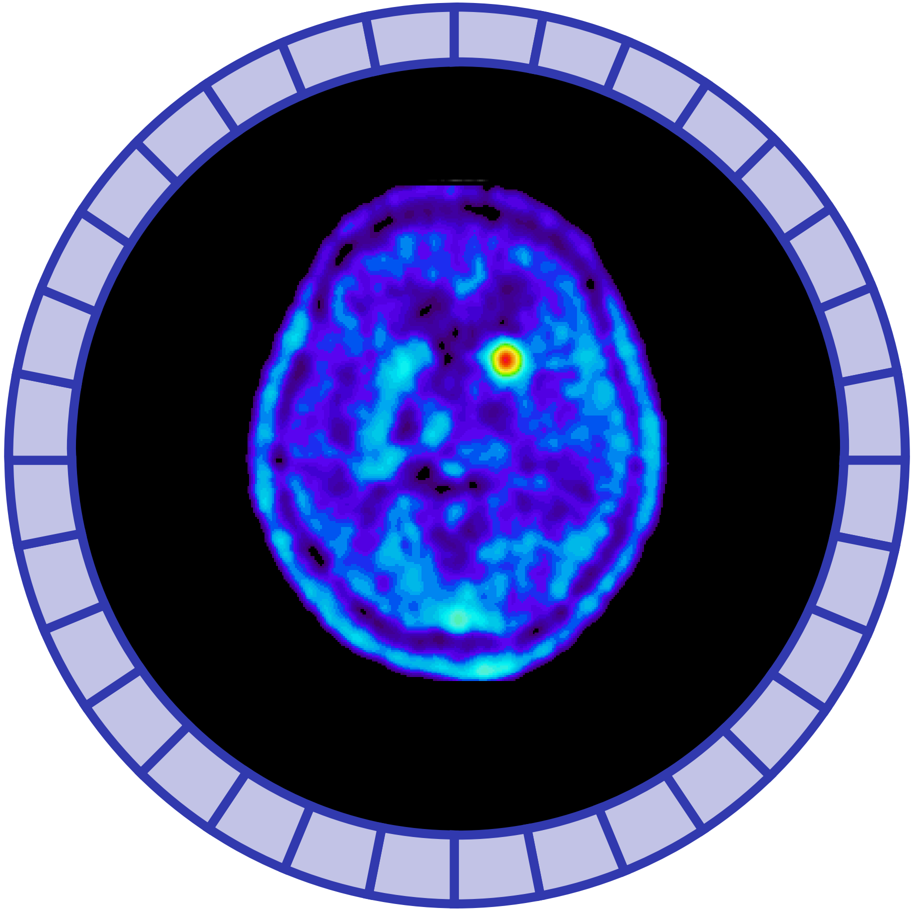
PYTORCH

European Funds Smart Growth
Republic of Poland
European Union European Regional Development Fund

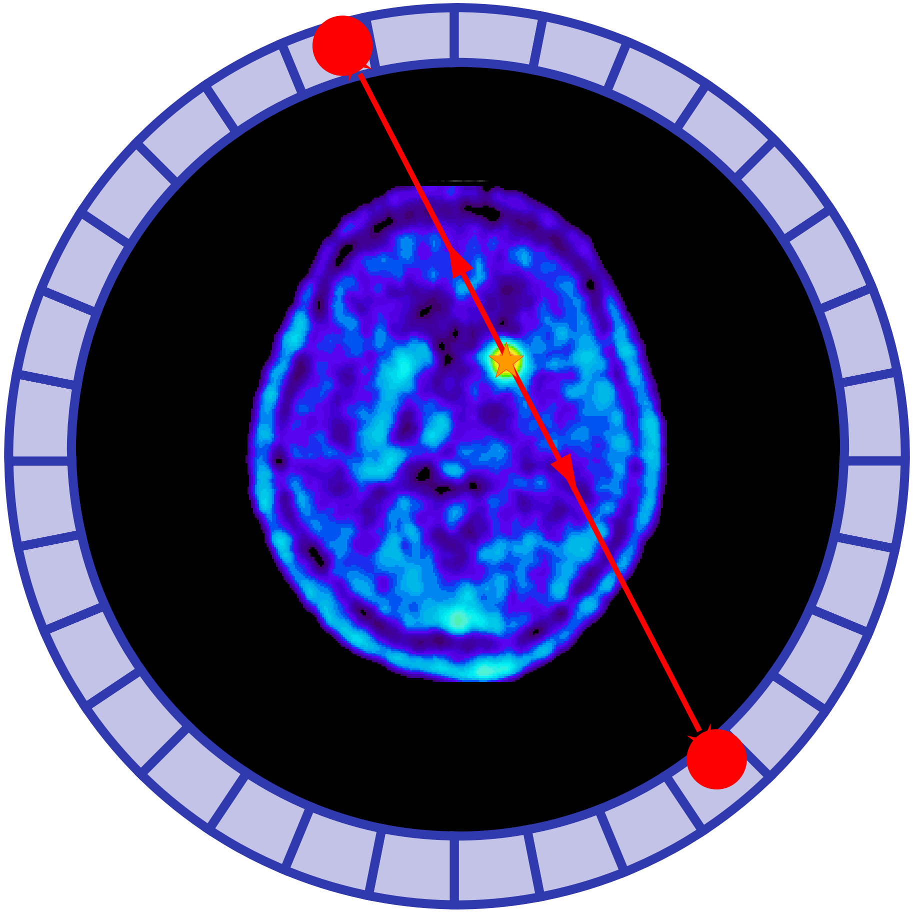
PET


ring of photon
detectors
11C
ν
decay by
positron
emission
γ
γ
β+
e-
γ photon
detection
γ photon
detection
annihilation

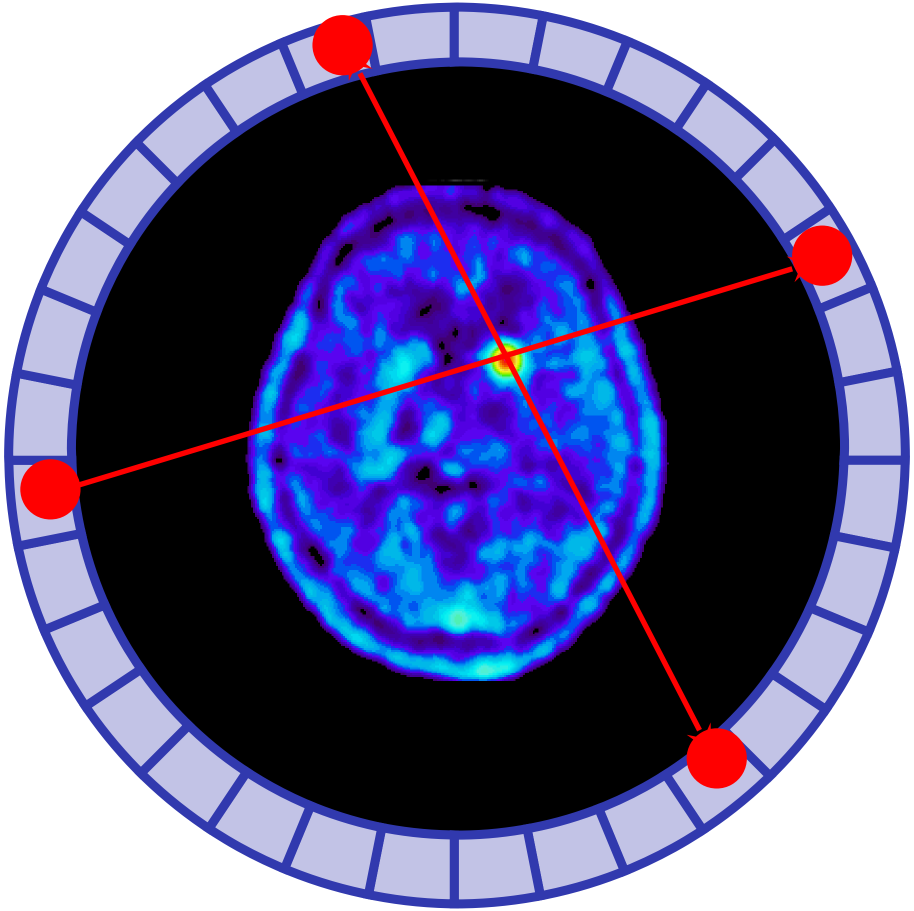
PET



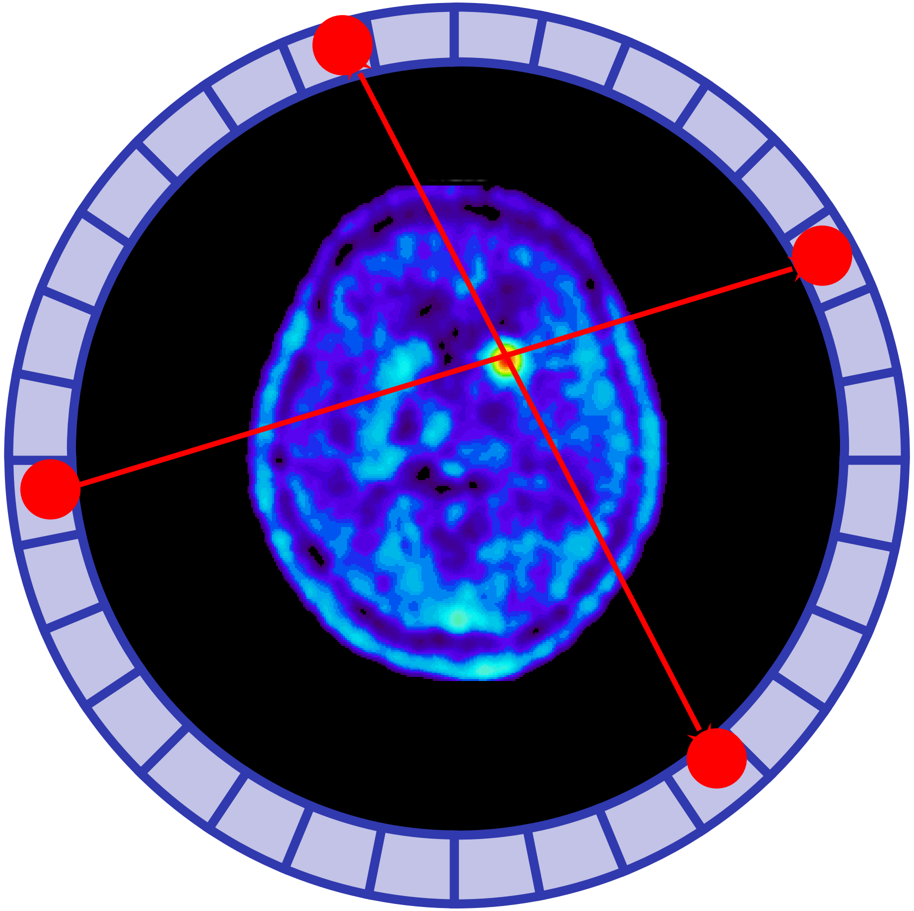
PET



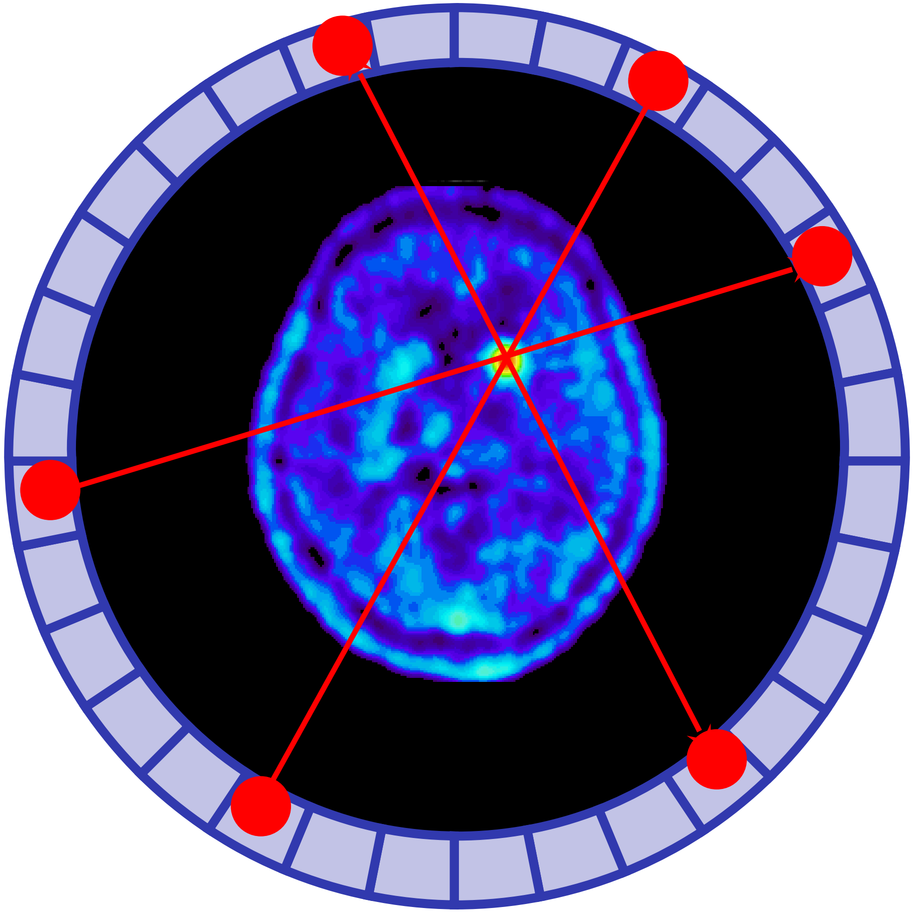
PET



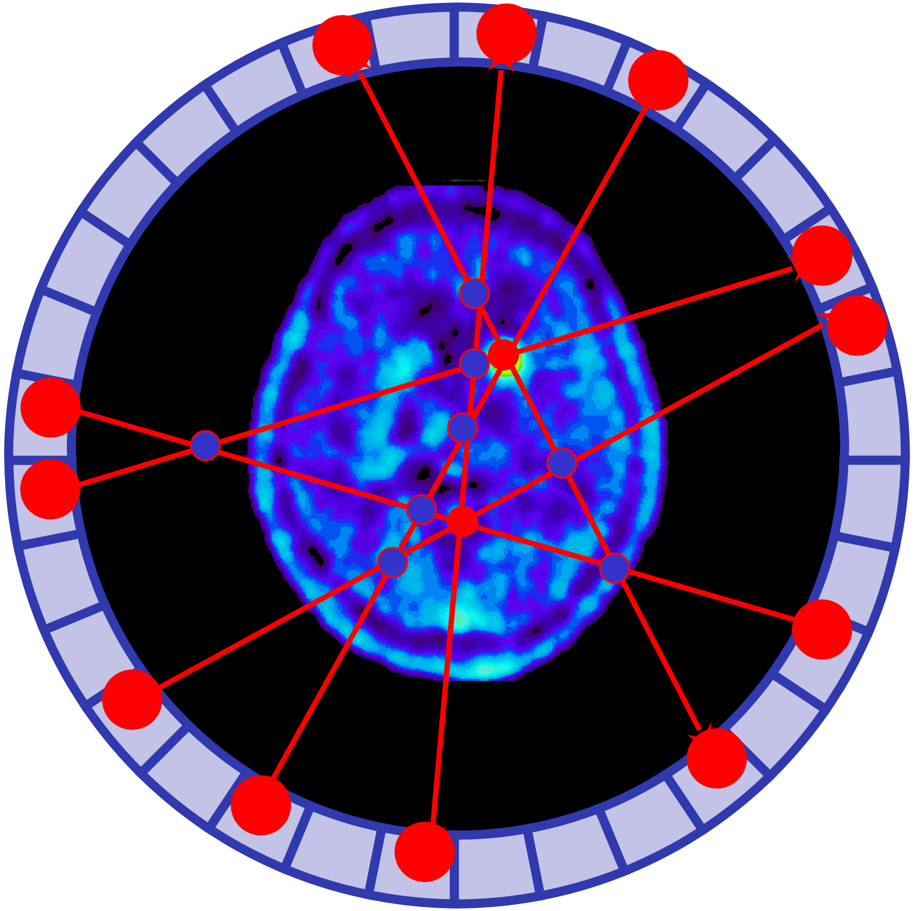
PET



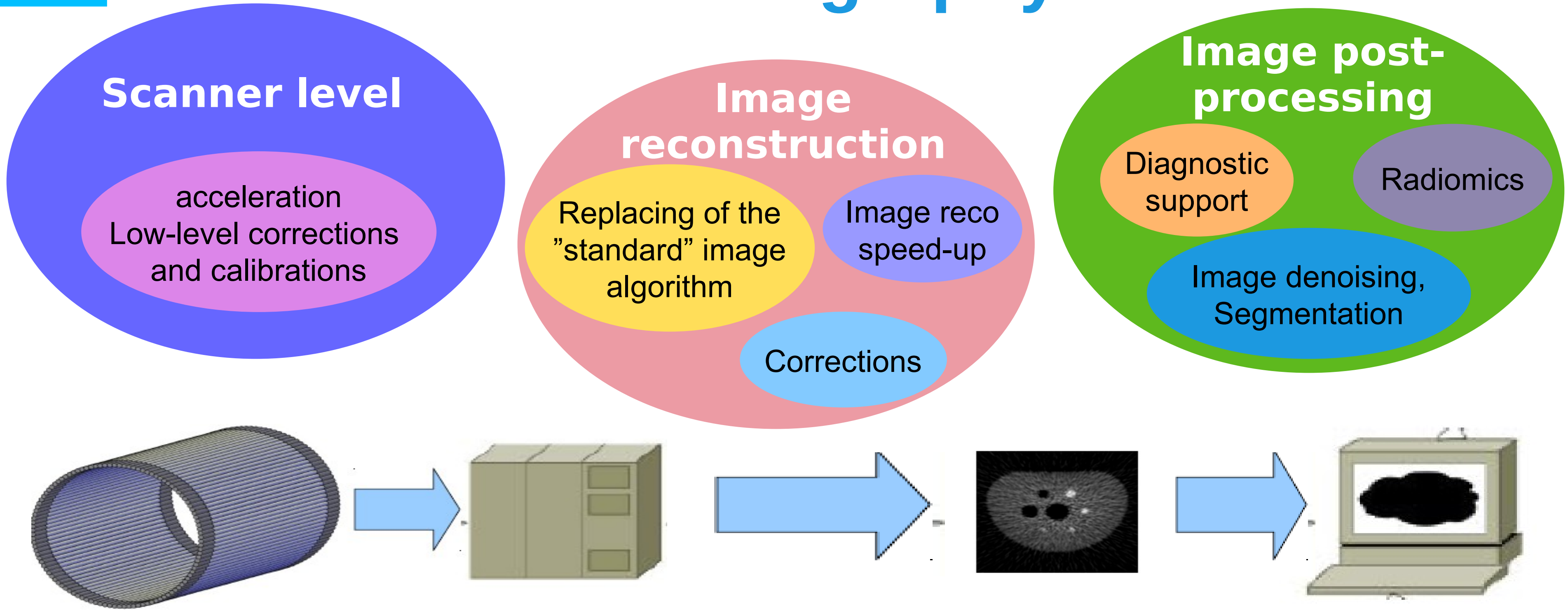
PET



PET



Artificial Intelligence methods in PET tomography



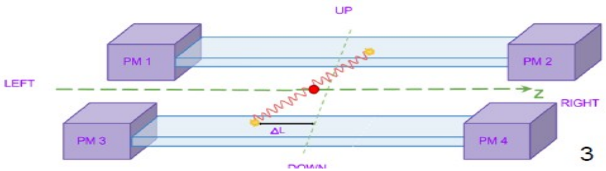
E.g. Using convolutional neural networks to estimate time-of-flight from PET detector waveforms
E. Berg and Simon R. Cherry Phys Med Biol . 2018 Jan 11;63(2):02LT01. doi: 10.1088/1361-6560/aa9dc5.

A Review of Deep-Learning-Based Approaches for Attenuation Correction in Positron Emission Tomography
J,S, Lee IEEE Transactions on Radiation and Plasma Medical Sciences 2020

Many papers about DL application for CT-free PET and image reco using DL

Cost-effective total body solution

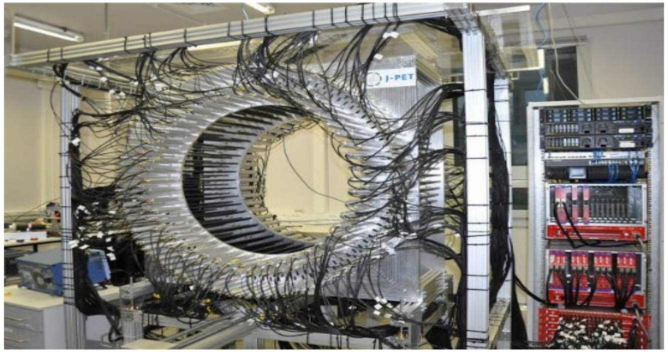
$$\Delta l = \frac{(t_2 - t_1)\cdot v}{2} \cong \frac{(t_2 - t_1)\cdot c}{4}$$



$$\Delta x = \frac{(t_l - t_r)\cdot c}{2} \Longrightarrow \Delta x = \frac{\Delta t}{2}\cdot c$$

First prototype

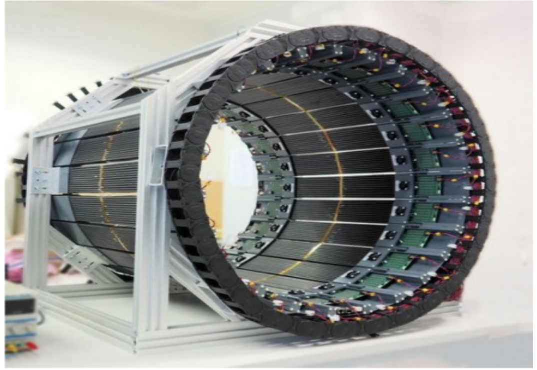
Acta Phys Pol. B 48 (2017) 1567



- 50 cm AFOV
- 192 plastic strips
- Readout → vacuum tube photomultipliers

Cost-effective total body solution

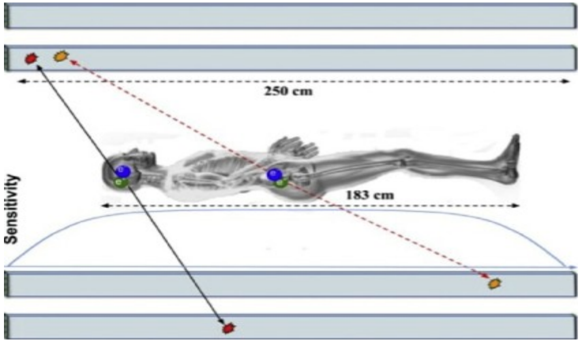
Modular J-PET



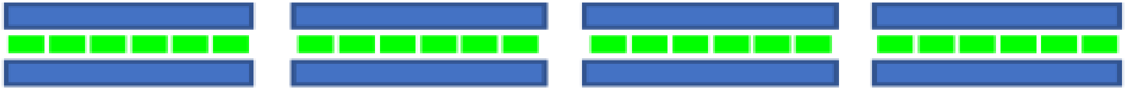
- 50 cm AFOV
- 24 modules x 13 strips
- Readout → silicon photomultipliers matrices

Total-body

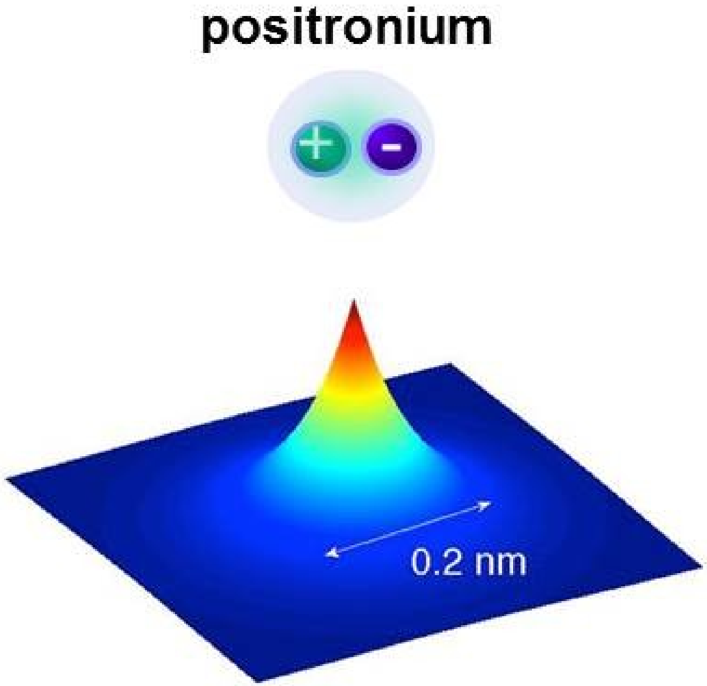
PET Clinics 15 (2020) 439
Phys. Med. Biol. 66 (2021) 175015

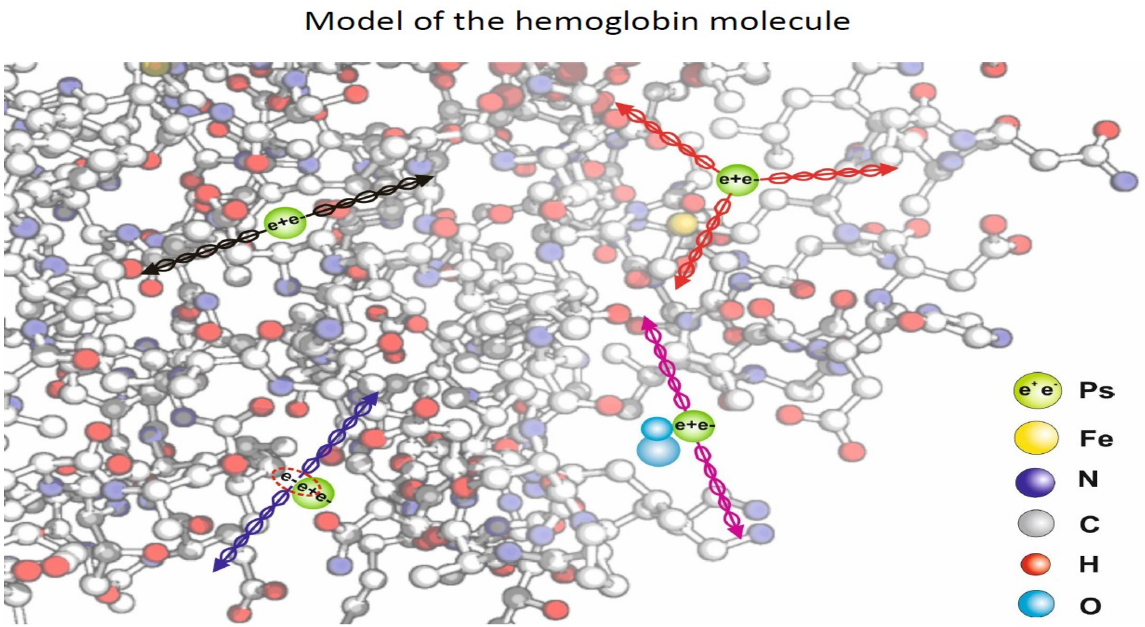


- 250 cm AFOV
- Additional layers of wavelength shifters → better axial resolution



Towards multiphoton/positronium (quantum) tomography





P. Moskal, B. Jasińska, E. Ł. Stępień, S. D. Bass
Nature Reviews Physics 1 (2019) 527-529

P. Moskal et al. Phys. Med. Biol. 64 (2019) 055017

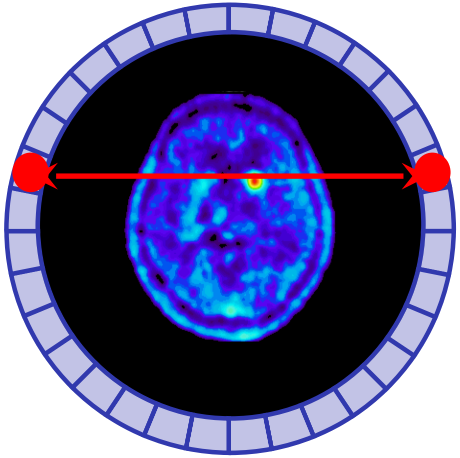
P. Moskal et al. EJNMMI Phys. 7 (2020) 44

P. Moskal, K. Dulski et al Science Advances 7 (2021) eabh4394

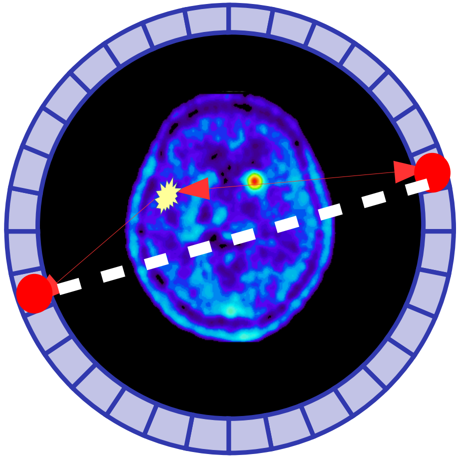
P. Moskal, A. Gajos et al. Nature Communications 12 (2021) 5658

Coincidence classification for total-body J-PET

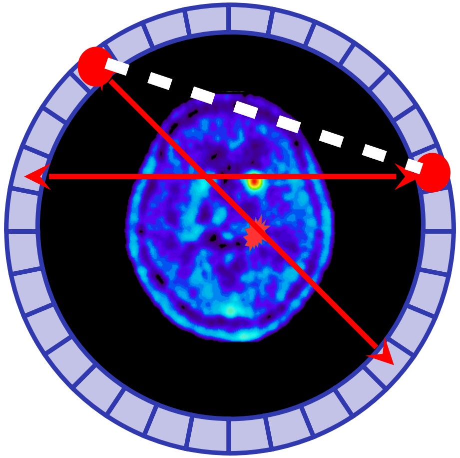
True



Scattered



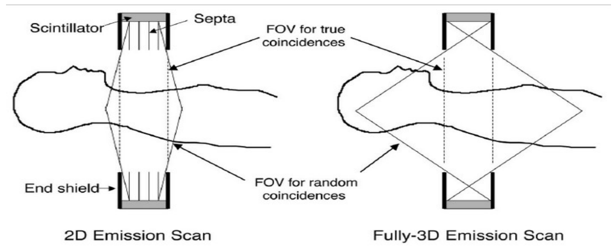
Accidental



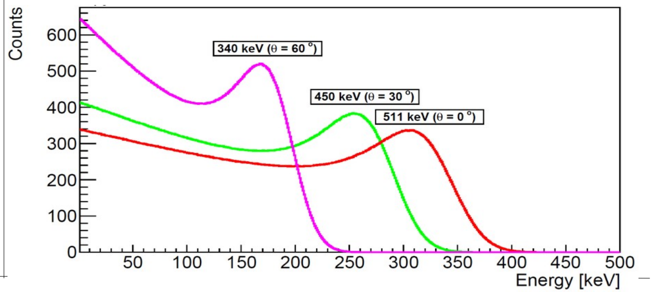
Coincidence classification for total-body J-PET

For total-body J-PET scanner we expect
higher background level from non-genuine coincidences

In J-PET



Multiple scattering in the phantom is not negligible



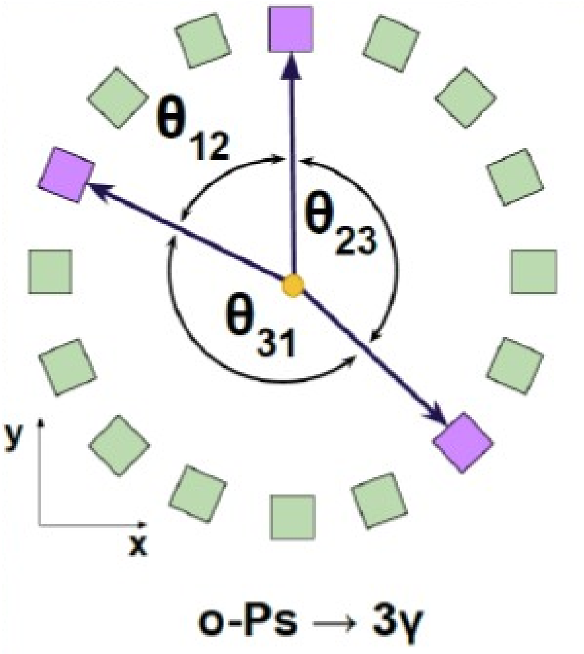
D. Brasse et al. J Nucl Med 2005; 46:859–867

Coincidence classification for total-body J-PET

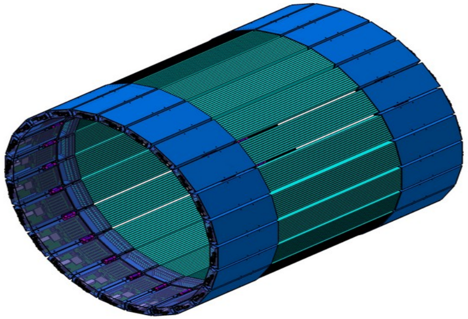
Situation much more complicated for multi-photon coincidences...

- More photons ➡ More combinations
- Less strictly defined geometry
- Photon energies have a distribution

Idea: apply ML techniques to reduce background (ACCIDENTAL, SCATTER)

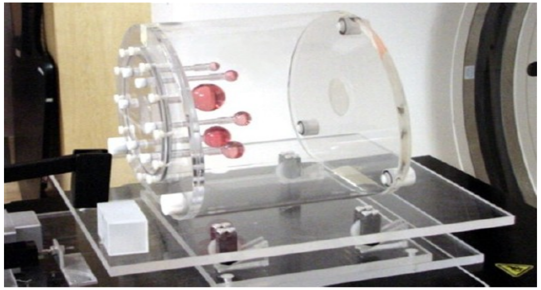


Training data generation



Modular J-PET

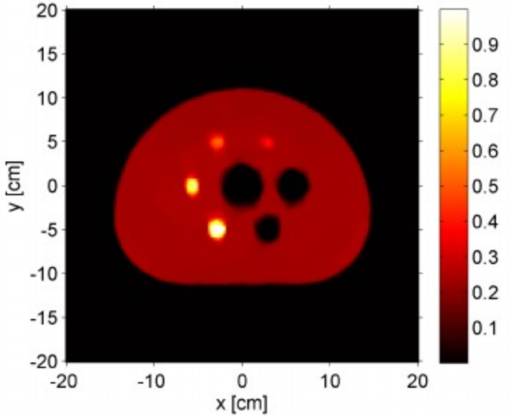
- 50 cm AFOV
- 24 modules x 13 strips
- 24 x 6 x 500 mm strips



NEMA IEC Phantom

- 4 hot spheres
- 2 cold spheres
- Activity - 59 Mbq
- acquisition time - 500 seconds
- contrast between hot and cold regions – 4:1

Monte Carlo Simulations

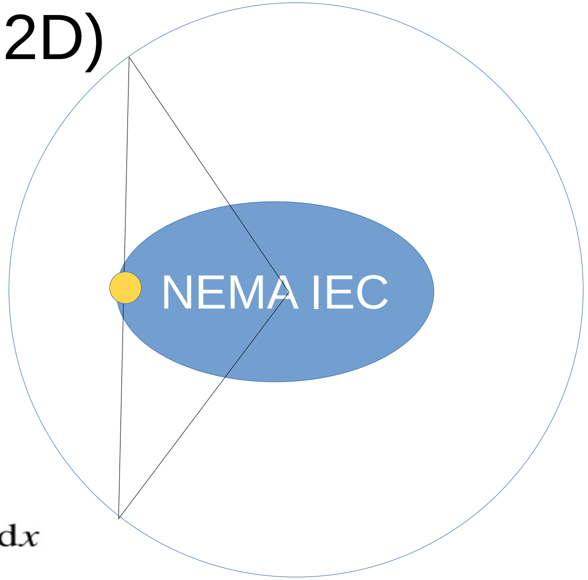


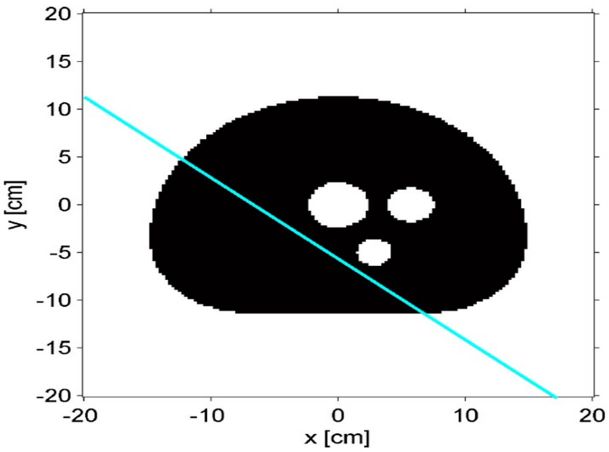
GATE MC Simulation

- 30M coincidences
- Phenomenological time, energy and positional resolution
- Geometry cuts → reduce accidental fraction

Features

1) 2D-angle with respect to the geom. center (deg2D)
2) time difference (dt)
3) LOR length (lorL)
4) energy difference $|E_1 - E_2|$ (eDiff)
5) energy sum (eSum)
6) attenuation coefficients (mu)



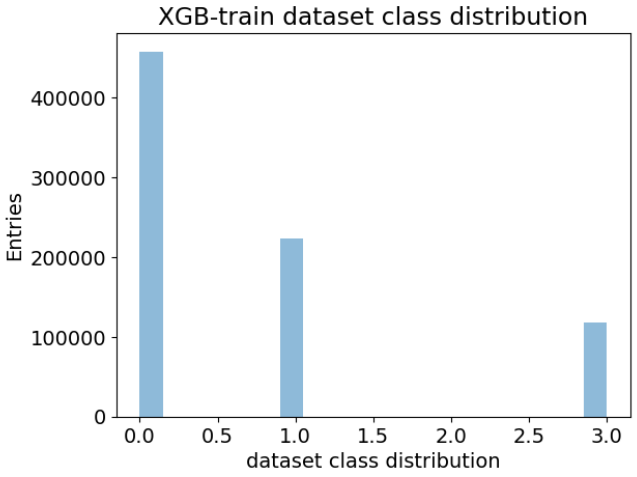


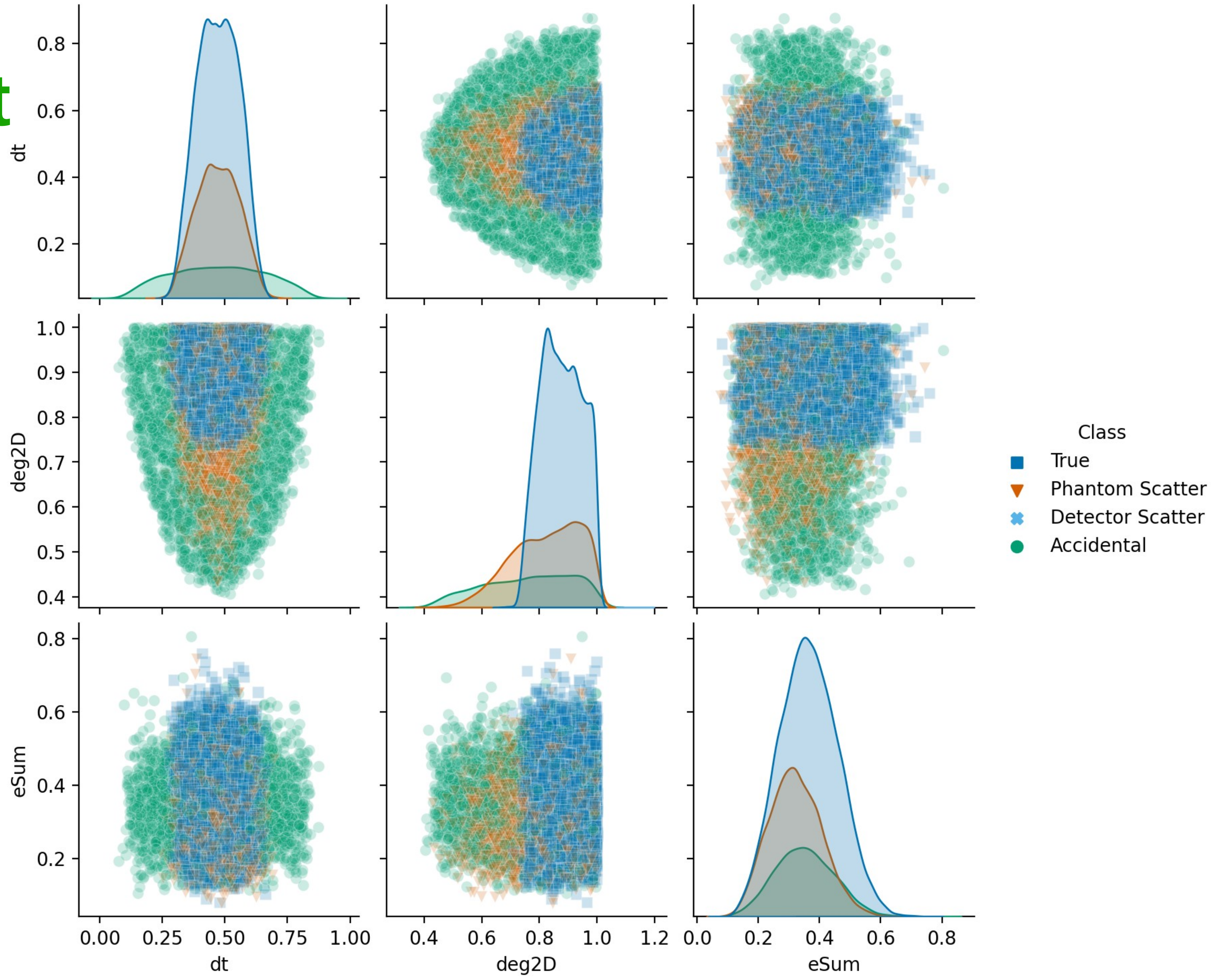
$$a = \int_{-\infty}^{+\infty} \exp(-\mu(x)x)\, \mathrm{d}x$$

$\mu = 0\ \mathrm{cm}^{-1}$ for air

$\mu = 0.096\ \mathrm{cm}^{-1}$ for water

Dataset





True: 57.3%

Classificators

- Feedforward Neural Network

Keras + SciKeras + scikit learn

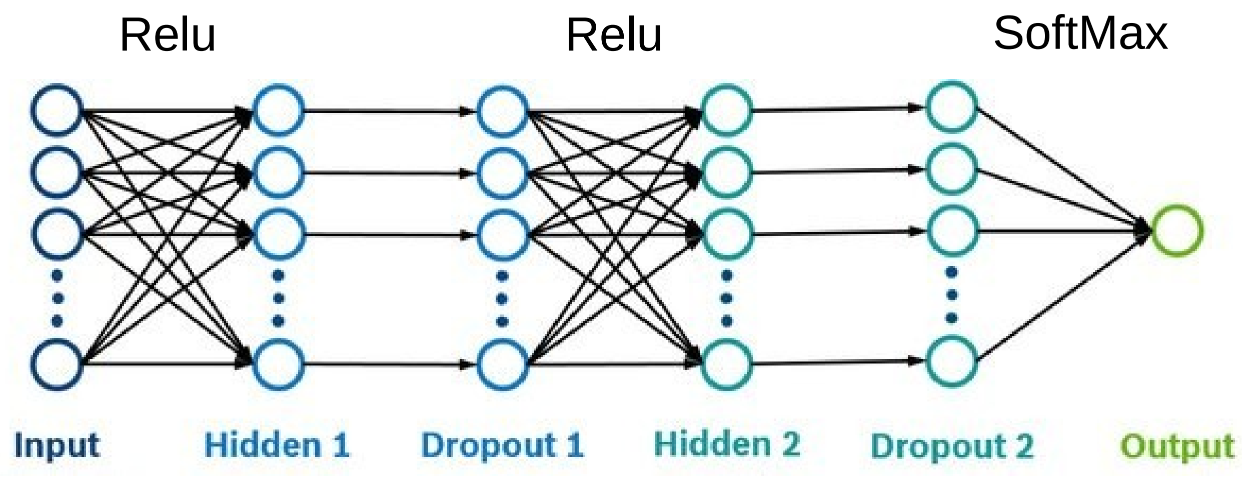
- ADABoost

scikit learn

- XGBoost

dmlc XGBoost + scikit learn

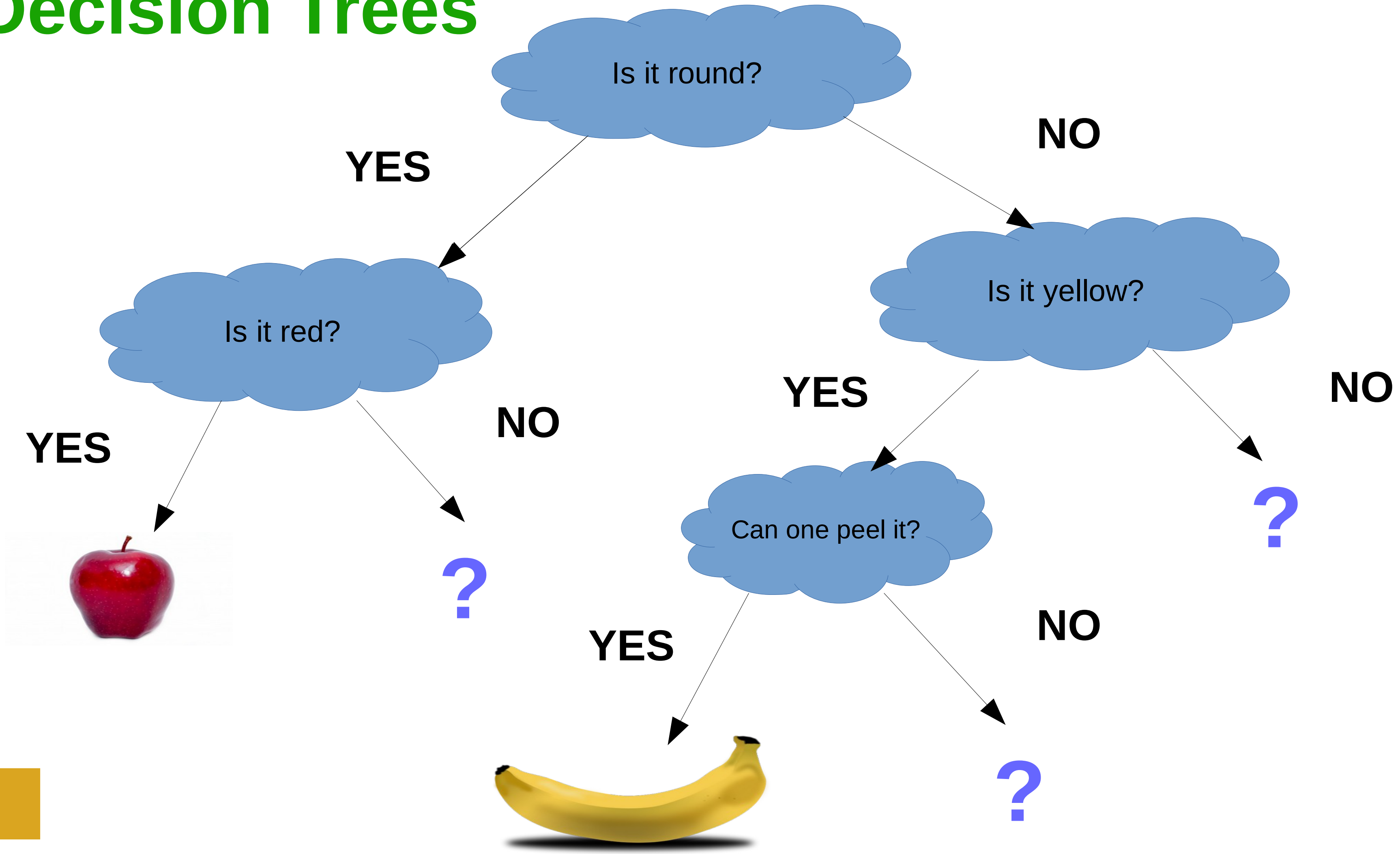
Feedforward Neural Network



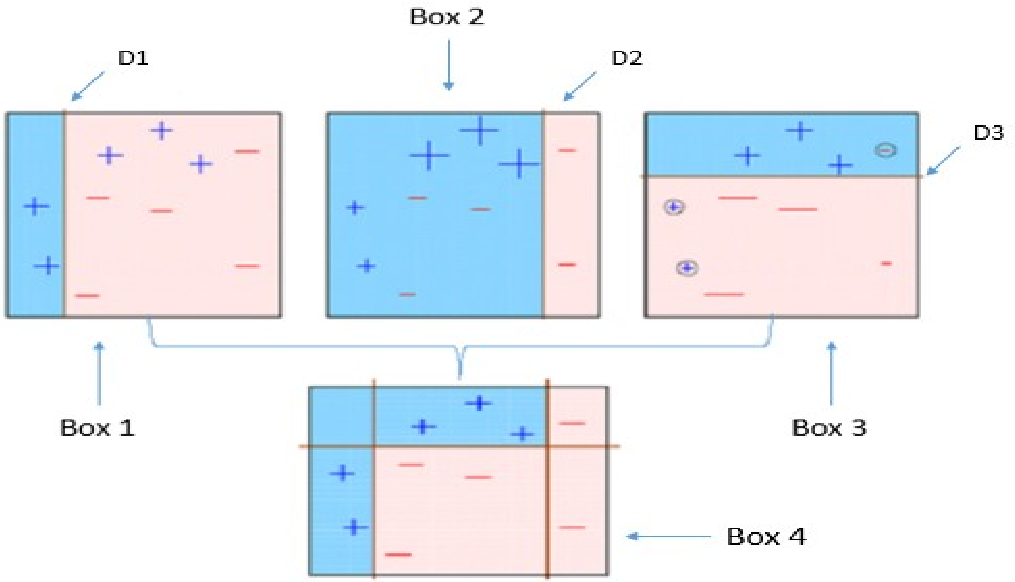
- Optimizer: Adam
- Loss: Categorical Cross Entropy
- Size and number of hidden layers optimized using Cross Validation
- Input data normalized to (0, 1) range

Figure from: Benjamin Volz et.al „A data-driven approach for pedestrian intention estimation”

Decision Trees



Boosting - AdaBoost



$$F_M(\vec{x}) = F_{M-1}(\vec{x}) + \lambda f_M(\vec{x})$$

$$\left.\frac{\partial L[y_i, F(\vec{x}_i)]}{\partial F(\vec{x}_i)}\right|_{F(\vec{x}_i) = F_{M-1}(\vec{x}_i)}$$

Gradient descent over functional space

Figures from https://www.analyticsvidhya.com/blog/2015/11/quick-introduction-boosting-algorithms-machine-learning/

https://github.com/QuentinSoubeyran/scikit-optimize/tree/unofficial

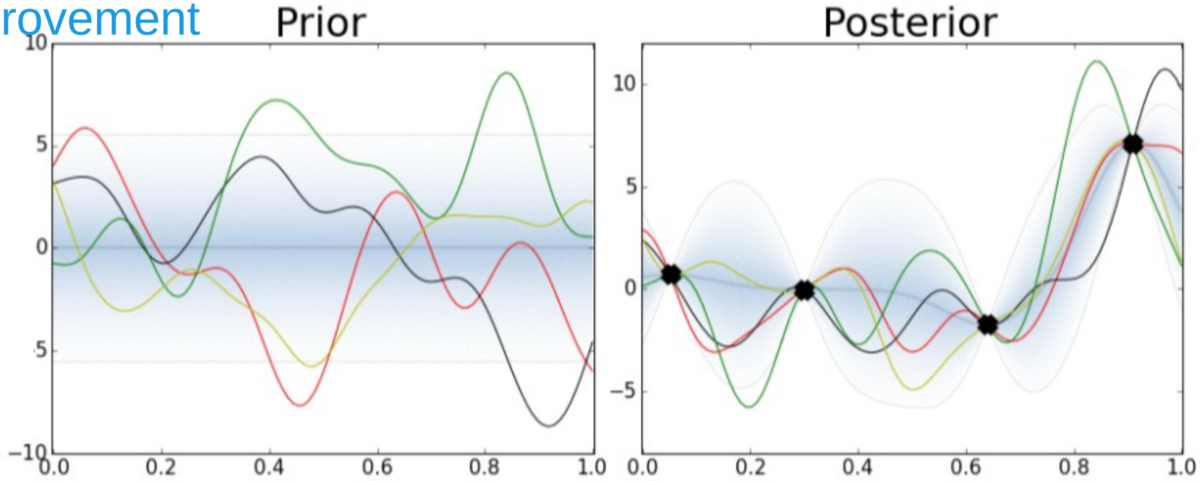
Bayesian optimization + cross validation

f – model accuracy on test sample

1) Choose some prior measure over the space of possible objectives f
2) Combine prior and the likelihood to get a posterior measure given some observations
3) From posterior decide where to take the next evaluation according to acquisitiotion function

- surrogate probability model – Gaussian process
- acquisition function – Expected Improvement
- f evaluation – best model from 3-fold cross validation

$$\alpha_{EI}(\mathbf{x};\theta,\mathcal{D}) = \int_y \max(0, y_{best} - y)p(y|\mathbf{x};\theta,\mathcal{D})dy$$

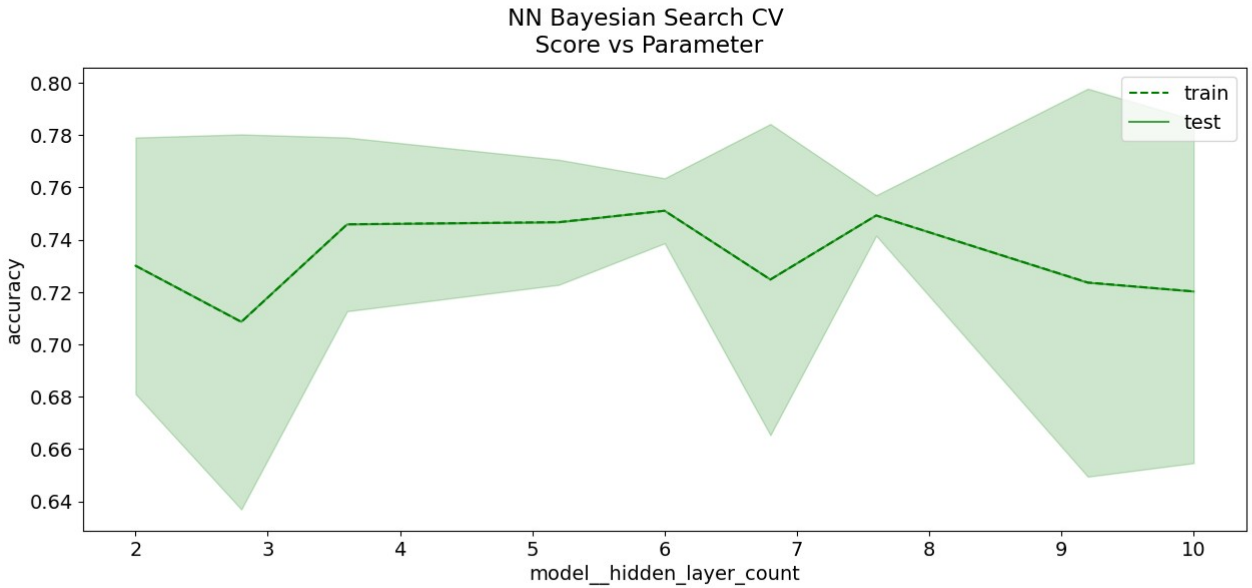


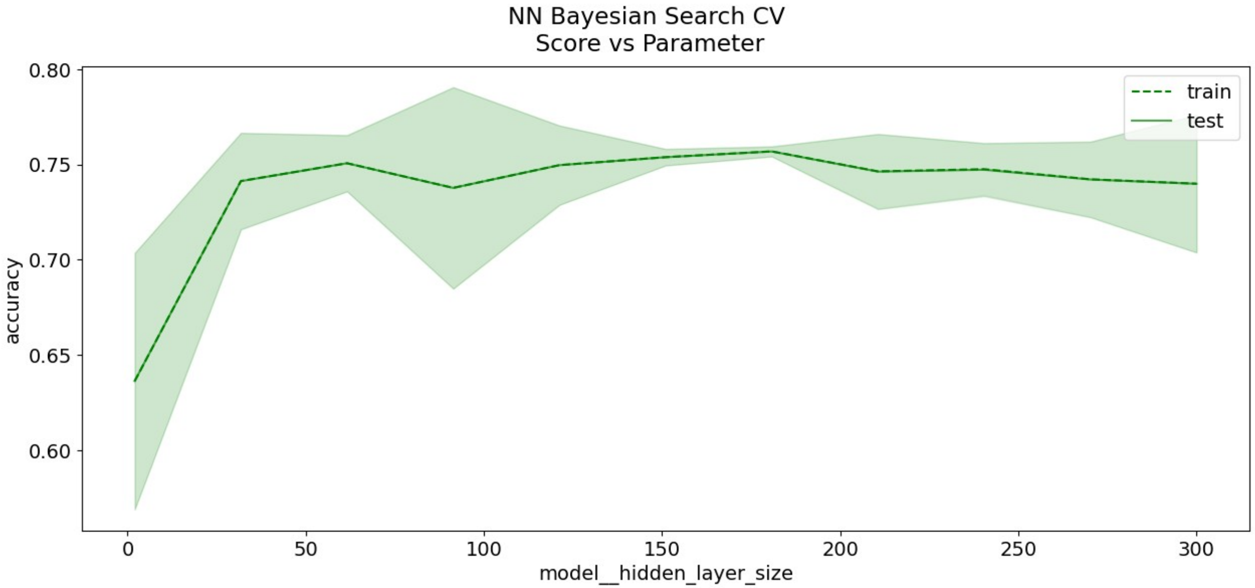
Figures from: Javier Gonzalez, „Introduction to Bayesian Optimization”

NN CV

Best model:

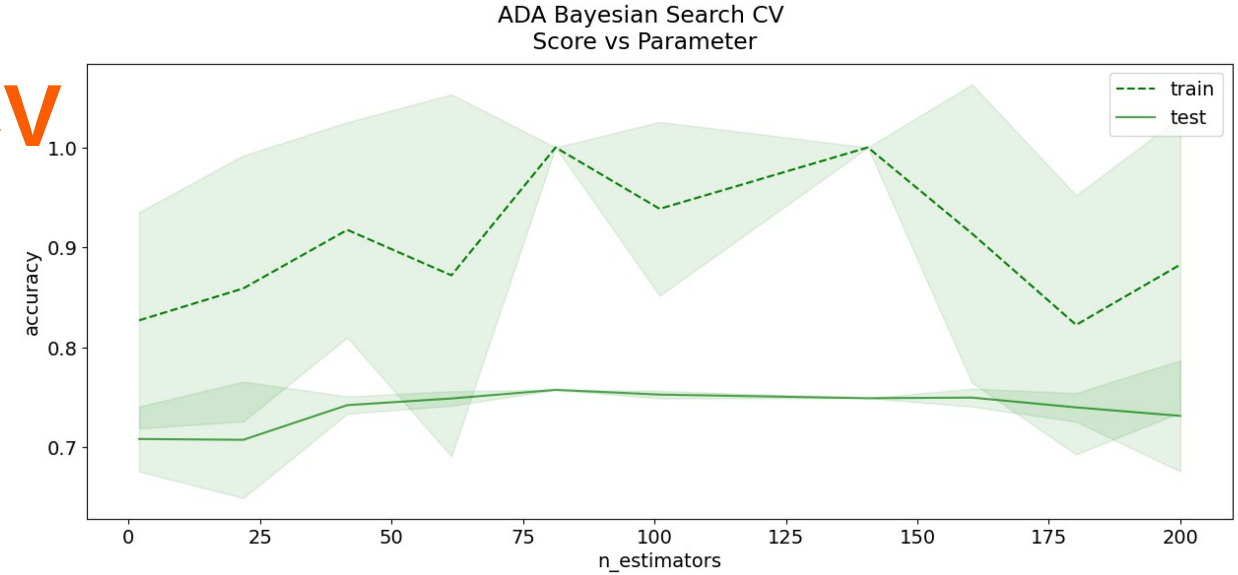
- batch size: 32
- dropout: 0.13
- hidden layers: 2
- layer size: 180
- learning rate: 0.001

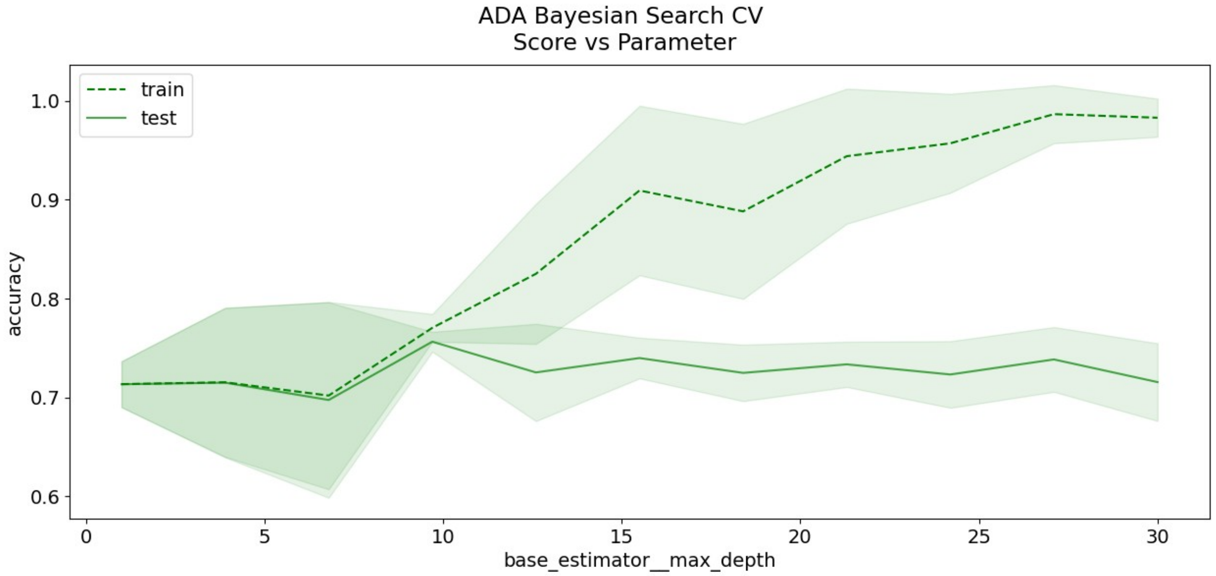




ADABoost CV

Best model:
- max_depth: 8
- learning_rate: 0.022
- n_estimators: 20

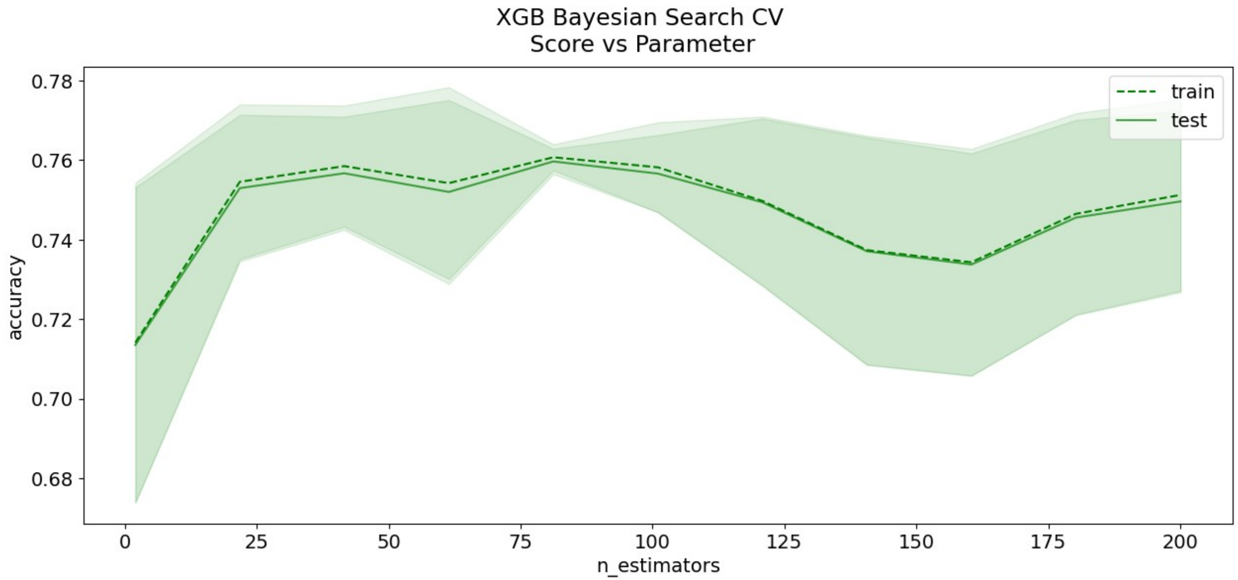


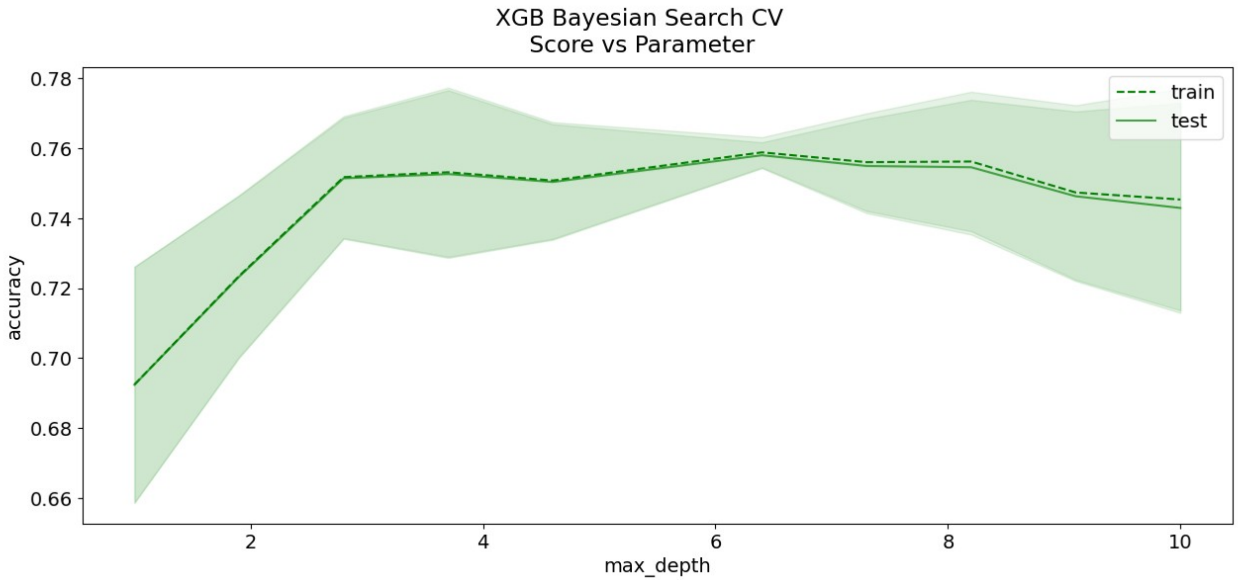


XGBoost CV

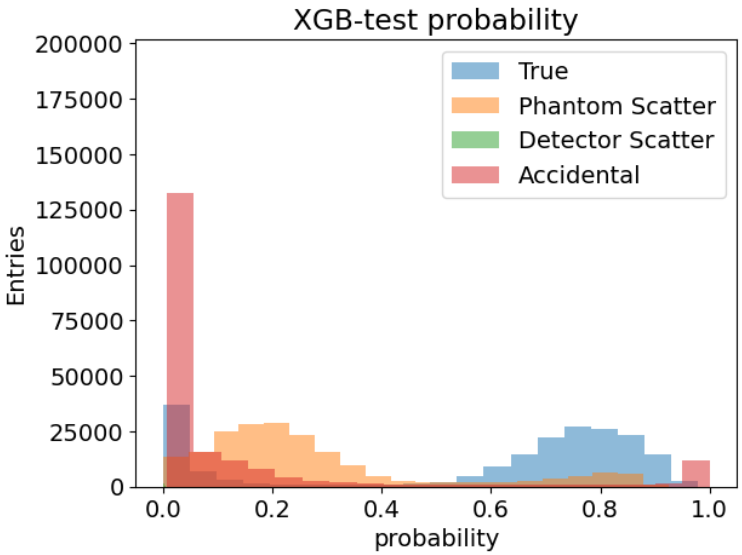
Best model:

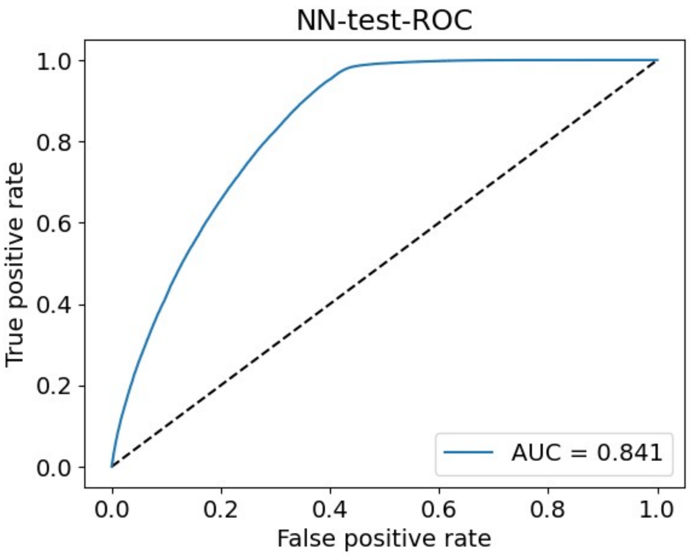
- colsample bytree: 0.96
- gamma: 0.12
- learning rate: 0.1
- max_depth: 10
- min_child_weight: 5
- estimators: 200
- reg_alpha: 6.23
- reg_lambda: 4.15
- subsample: 0.97

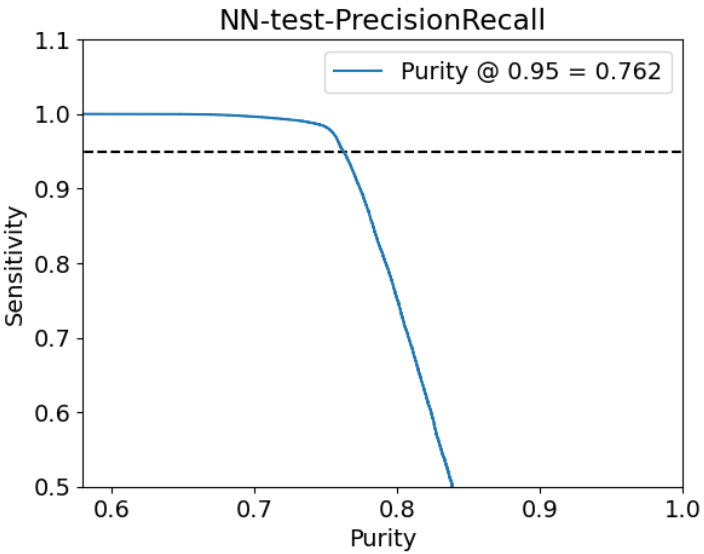


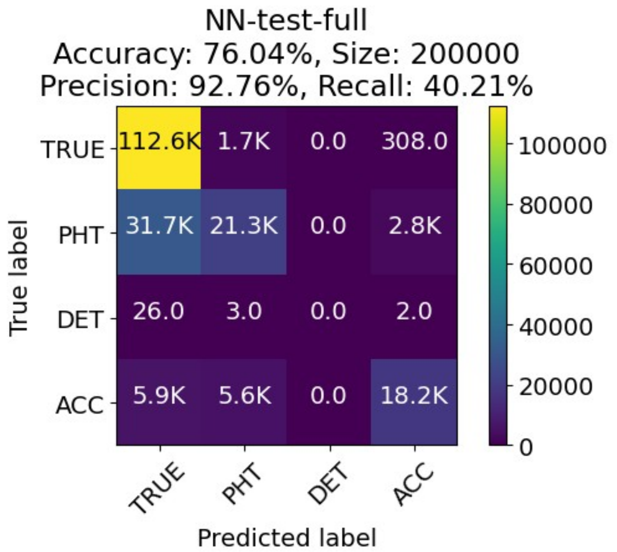


NN

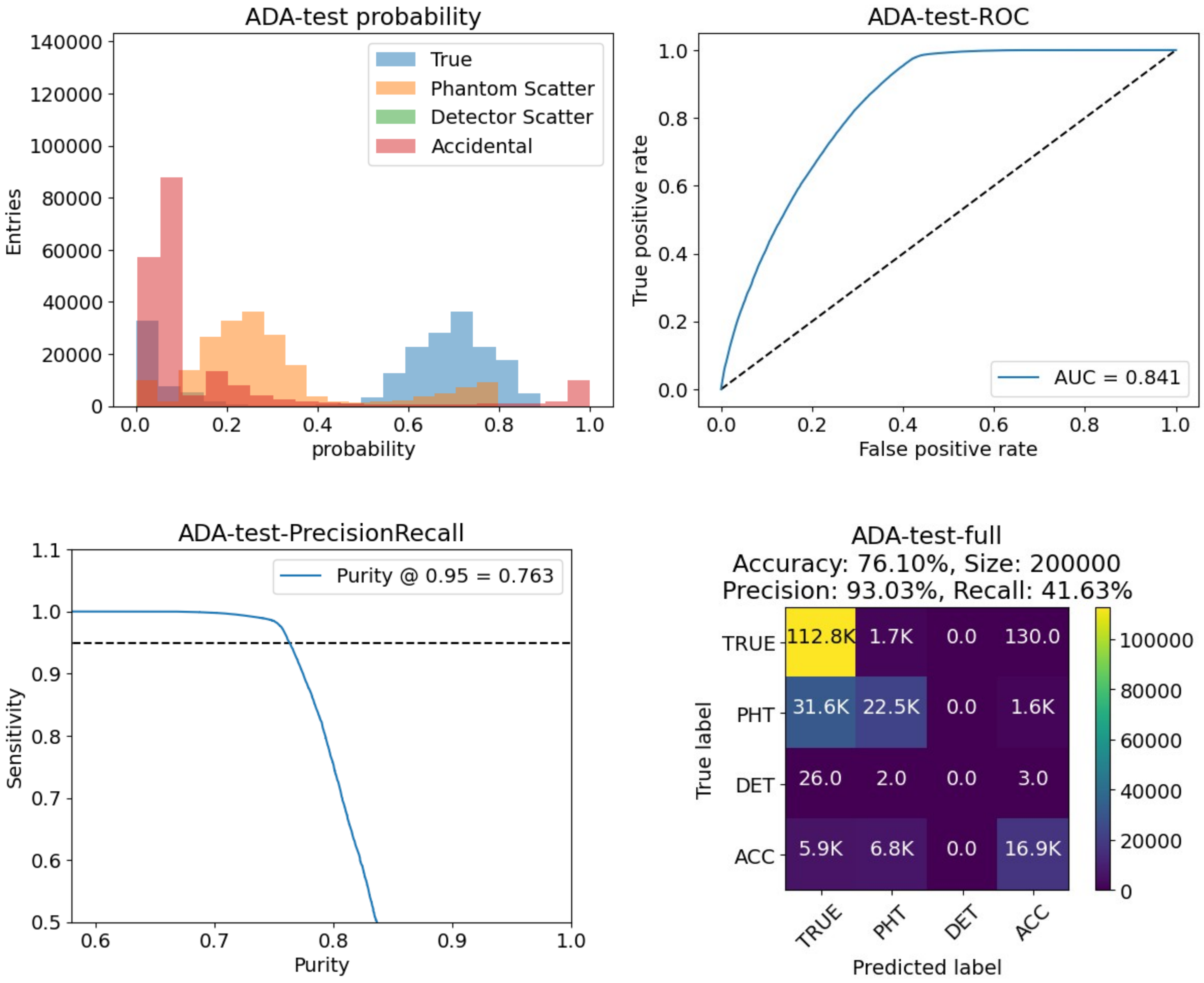

XGB-test probability
True
Phantom Scatter
Detector Scatter
Accidental
Entries
probability
0
25000
50000
75000
100000
125000
150000
175000
200000
0.0
0.2
0.4
0.6
0.8
1.0


NN-test-ROC
AUC = 0.841
True positive rate
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0

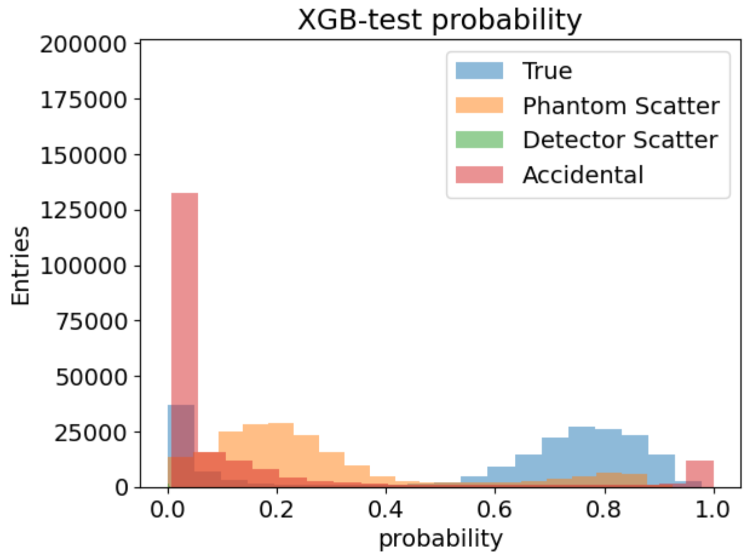

NN-test-PrecisionRecall
Purity @ 0.95 = 0.762
Sensitivity
Purity
0.5
0.6
0.7
0.8
0.9
1.0
1.1


NN-test-full
Accuracy: 76.04%, Size: 200000
Precision: 92.76%, Recall: 40.21%
TRUE
PHT
DET
ACC
112.6K
1.7K
0.0
308.0
31.7K
21.3K
0.0
2.8K
26.0
3.0
0.0
2.0
5.9K
5.6K
0.0
18.2K
True label
Predicted label
0
20000
40000
60000
80000
100000

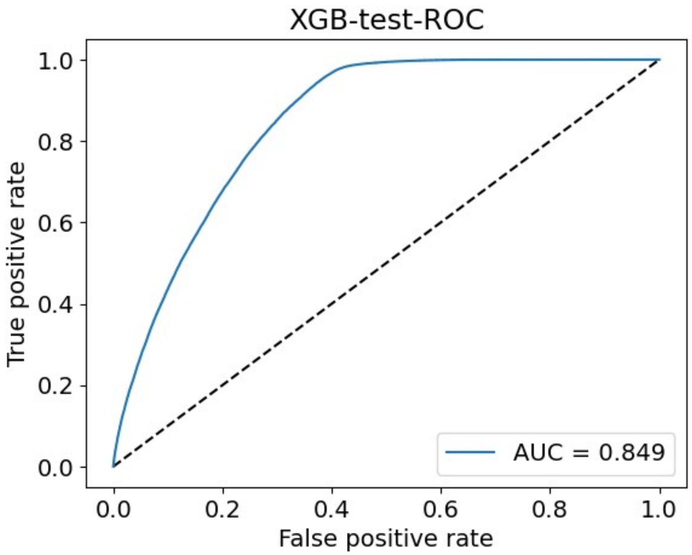
ADABoost


ADA-test probability
Entries
probability
True
Phantom Scatter
Detector Scatter
Accidental
ADA-test-ROC
True positive rate
False positive rate
AUC = 0.841
ADA-test-PrecisionRecall
Sensitivity
Purity
Purity @ 0.95 = 0.763
ADA-test-full
Accuracy: 76.10%, Size: 200000
Precision: 93.03%, Recall: 41.63%
True label
Predicted label
TRUE
PHT
DET
ACC
112.8K
1.7K
0.0
130.0
31.6K
22.5K
0.0
1.6K
26.0
2.0
0.0
3.0
5.9K
6.8K
0.0
16.9K

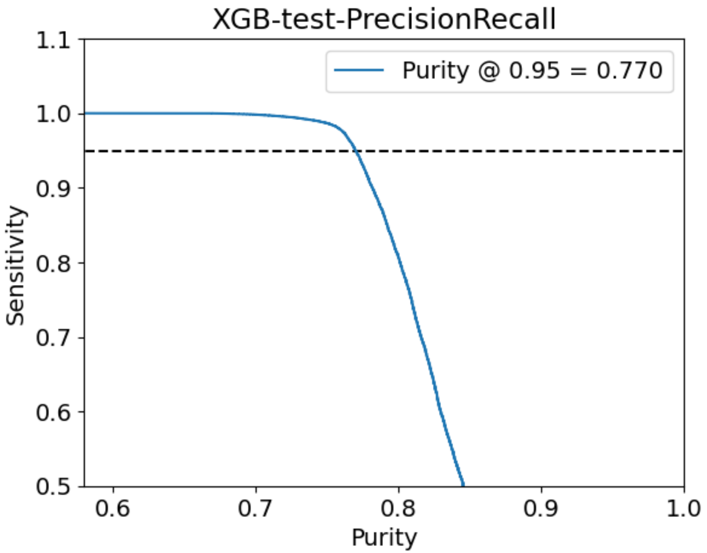
XGBoost



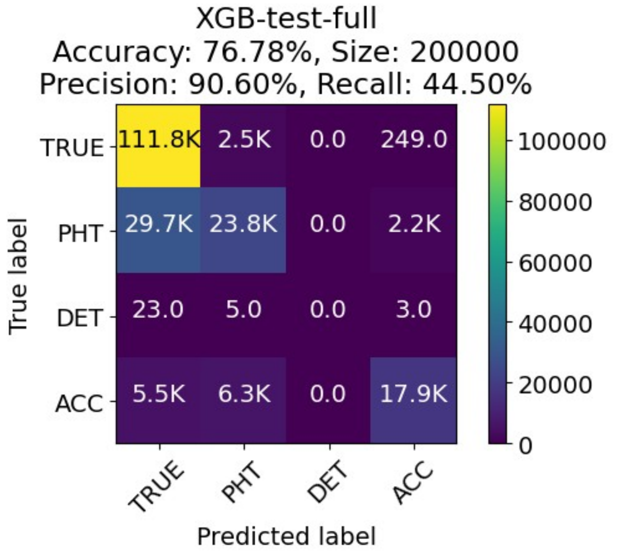
XGB-test probability
True
Phantom Scatter
Detector Scatter
Accidental
Entries
probability
200000
175000
150000
125000
100000
75000
50000
25000
0
0.0
0.2
0.4
0.6
0.8
1.0



XGB-test-ROC
AUC = 0.849
True positive rate
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0

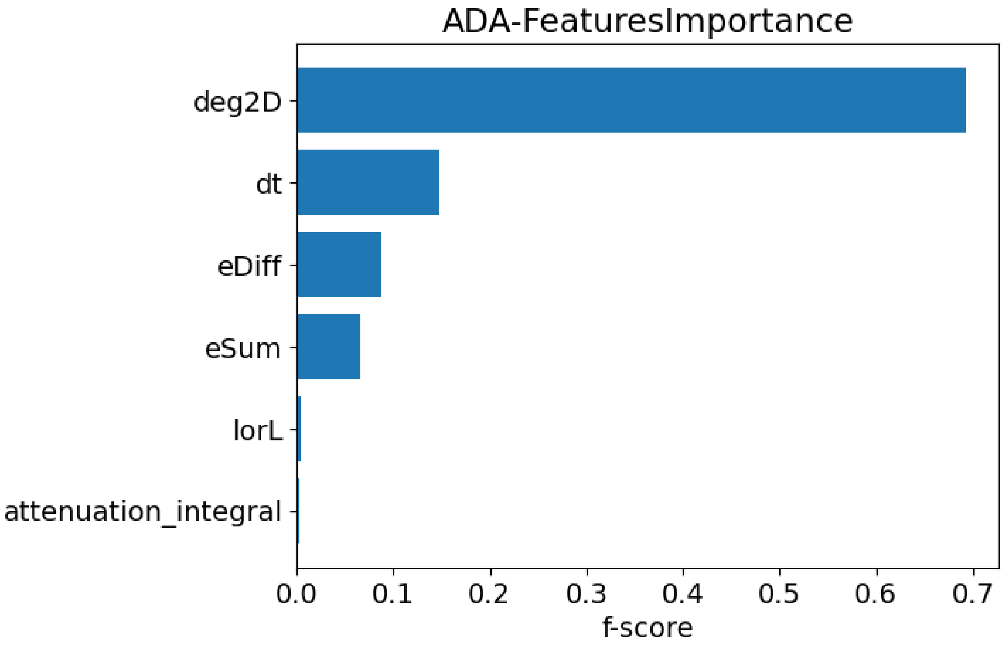


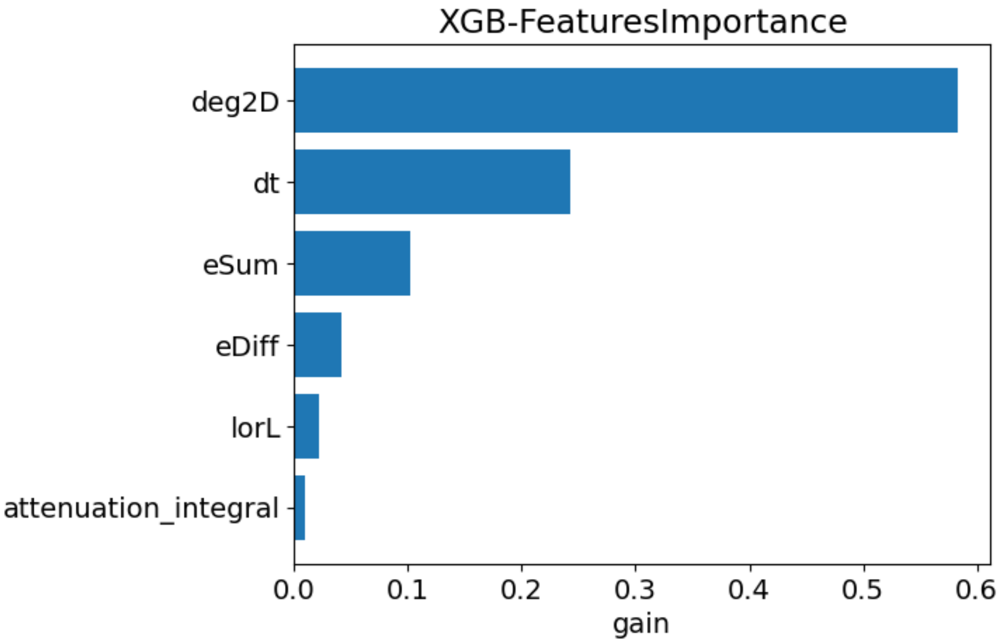
XGB-test-PrecisionRecall
Purity @ 0.95 = 0.770
Sensitivity
Purity
1.1
1.0
0.9
0.8
0.7
0.6
0.5



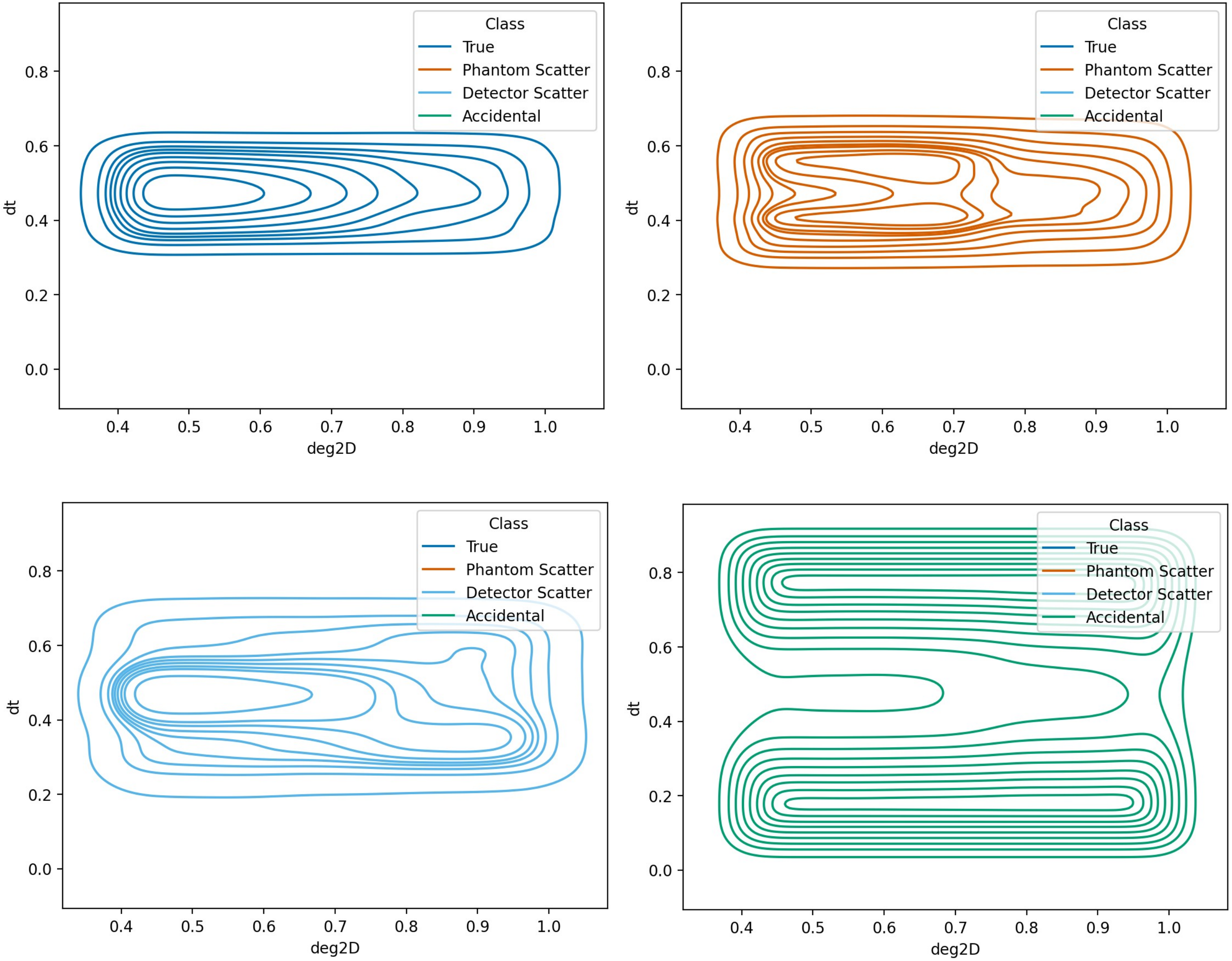
XGB-test-full
Accuracy: 76.78%, Size: 200000
Precision: 90.60%, Recall: 44.50%
TRUE
PHT
DET
ACC
111.8K
2.5K
0.0
249.0
29.7K
23.8K
0.0
2.2K
23.0
5.0
0.0
3.0
5.5K
6.3K
0.0
17.9K
True label
Predicted label
100000
80000
60000
40000
20000
0

Feature importance





XGBoost



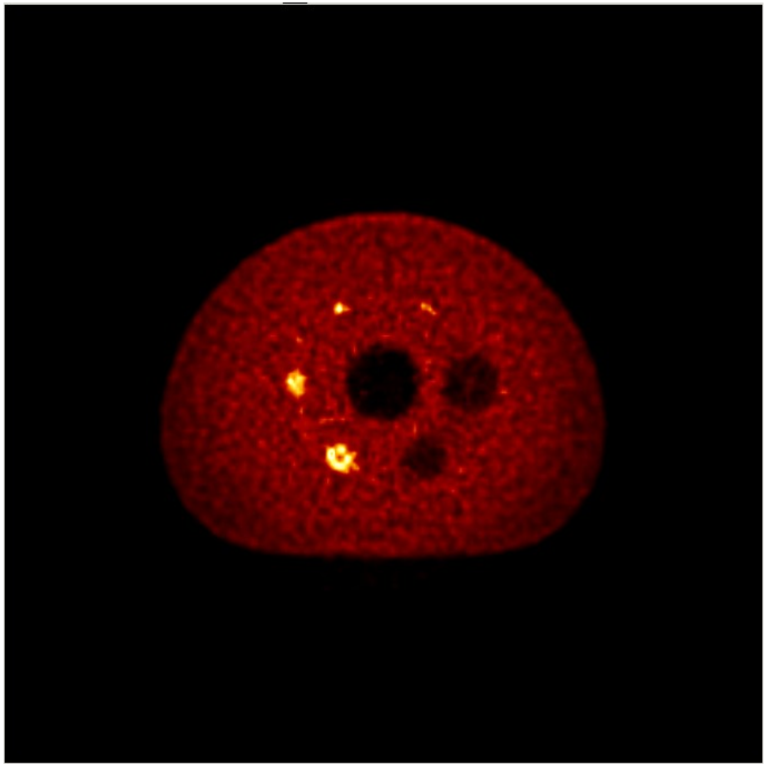
Class
True
Phantom Scatter
Detector Scatter
Accidental
dt
deg2D
0.0
0.2
0.4
0.6
0.8
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Image reconstruction

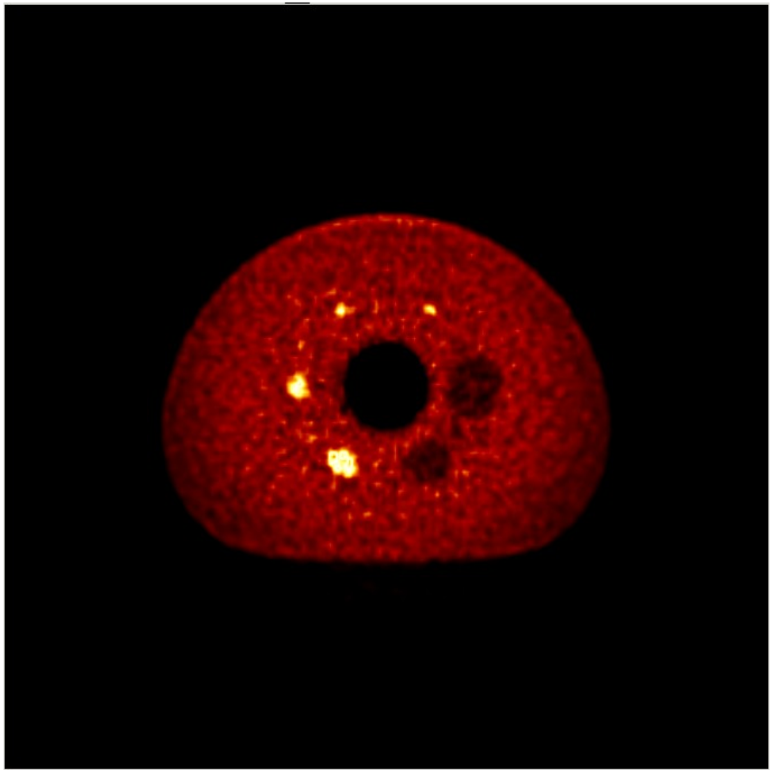
MLEM (CASTOR)

30 iterations

Without filtration



XGBoost Filtered



Summary

Method	Precision
Base line	57.3%
NN	76.2%
ADABoost	76.3%
XGBoost	77.0%

Summary

- Verification of ML applicability for PET coincidence classification
- Tested full chain with bayesian optimization
- Comparison of NN, ADABoost and XGBoost performance

- Information accessible in the two photon data
  does not allow for clear class separation.

  Further work needed.

Outlook

- Verify the model invariance wrt used phantom
  - test on different phantoms (e.g. X-CAT)
- Apply the method to multiphoton data
  - harder classification problem
  - however larger feature space
- Validation with real data