

Report on KOTO EMCal Study

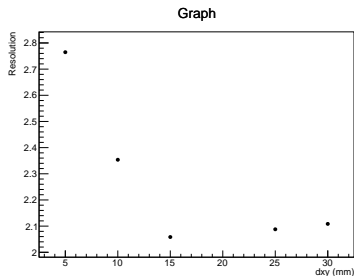
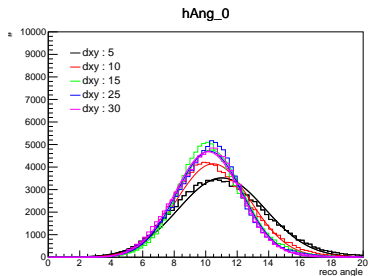
Junlee Kim

February 2, 2021

Contents

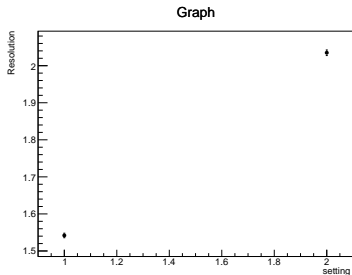
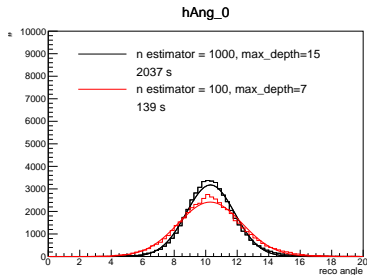
- ▶ Reconstruction as a function of the detector width
- ▶ Study on XGBOOST

dependence on the detector width



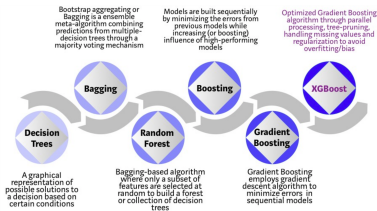
- ▶ Training sample: random generation for polar(0–50 deg) and azimuthal angle(0–360 deg) with 100k events
- ▶ Test samples: 100k fixed $\theta = 10^\circ$ events.
 - ▶ dxy is set in geant4 simulation.

Enhanced machine learning performance



- ▶ Training sample: random generation for polar(0–50 deg) and azimuthal angle(0–360 deg) with 100k events
- ▶ Test samples: 50k fixed $\theta = 10^\circ$ events.
 - ▶ Time for training governs CPU time (99%).
 - ▶ Time for test less than 10 s

XGBOOST

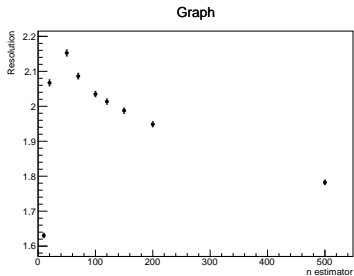
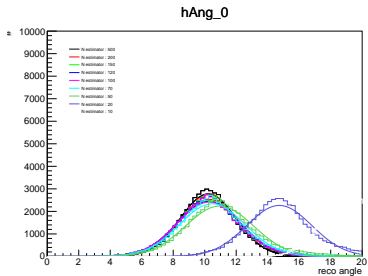


- ▶ <https://towardsdatascience.com/https-medium-com-vishalmorde-xgboost-algorithm-long-sh~:text=What%20is%20XGBoost%3F,all%20other%20algorithms%20or%20frameworks>.

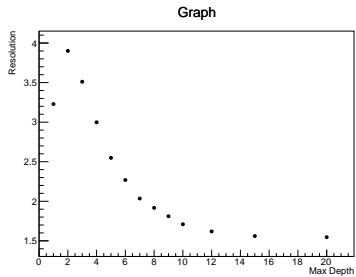
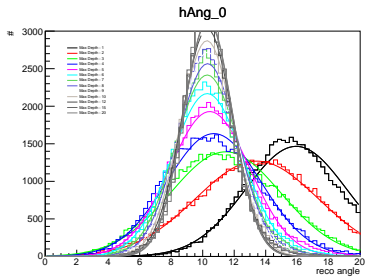
Machine learning parameters

- ▶ `n_estimators`: the number of decision trees
- ▶ `max_depth`: the number of stages of decision tree group
- ▶ The maximum depth of decision trees
 - ▶ A number of trees would be necessary to build a deep decision tree
 - ▶ Correlated.
- ▶ `learning_rate`: step size for regression
- ▶ `gamma`: minimum loss function requirement
- ▶ `subsample`: Rate of samples used for estimation.

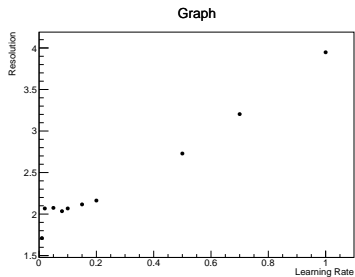
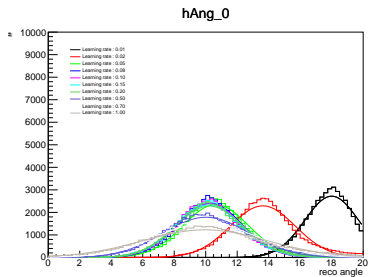
n_estimators



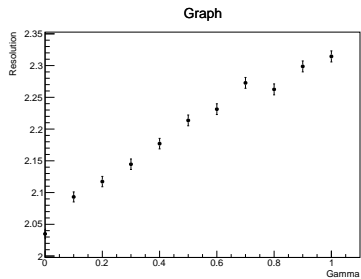
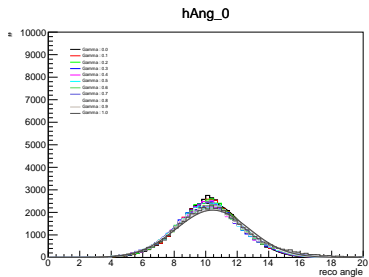
max_depth



learning_rate

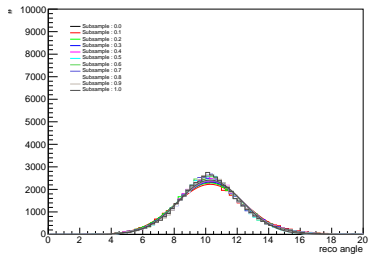


gamma

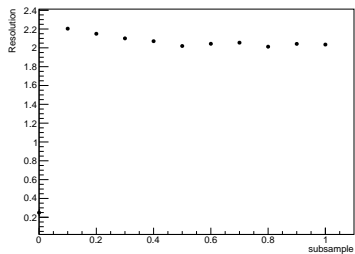


subsample

hAng_0



Graph



status

- ▶ Saving training results will be tested.
- ▶ Further investigation to optimize training?
- ▶ x/z and y/z will be estimated separately.
 - ▶ φ estimation
- ▶ Interpolation of origin from COE will be tested?
- ▶ Other suggestions?