Application of Machine Learning to Particle Identification at the BESIII experiment

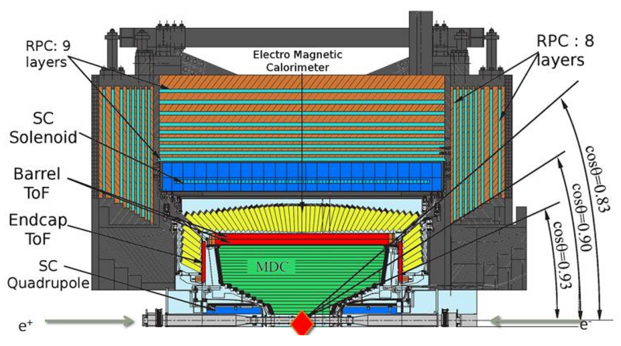
E-mail: chenzhengyuan@ihep.ac.cn

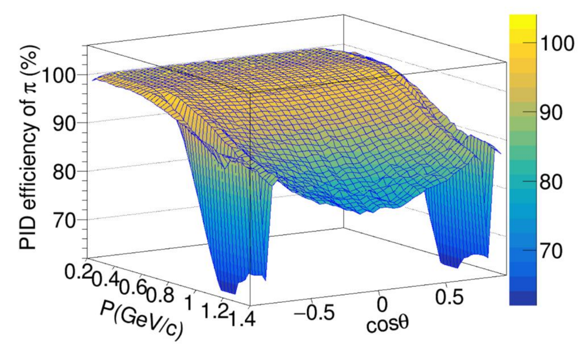
Zhengyuan Chen[1,2], Shengsen Sun[1,2], Huaimin Liu[1], Gang Li[1], Xiaobin Ji[12], Guang Zhao[1], Xingtao Huang[3], Tong Chen[12], Fang Liu[1], Xiaoling Li[3], Teng Li[3], Shuopin Wen[1], Chunxiu Liu[1], Wenxin Fang[1], Liangliang Wang[1]

[1]Institute of High Energy Physics Chinese Academy of Sciences
[2]Chinese Academy of Sciences
[3]Shandong university

Introduction

- Particle identification(PID) is an important ingredient to particle physics experiments
- Efficient hadronic identification usually requires the combination of variables from several dedicated detectors
- Machine learning algorithms have been developing to use computational methods to "learn" information directly from data
- With high statistic and high purity data samples, BESIII experiment offers an unique opportunity to take advantages of machine learning techniques to obtain a better performance of particle identification

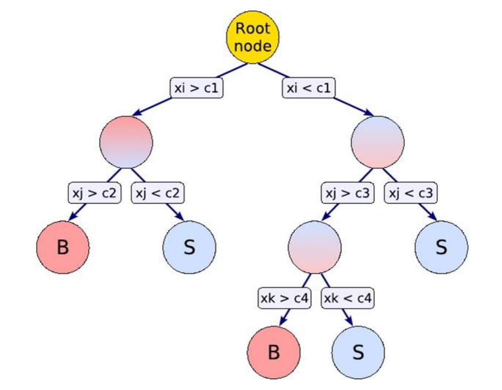




Current hadronic identification in BESIII

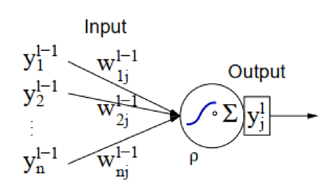
Machine learning models

Gradient boosted decision trees(BDTG)

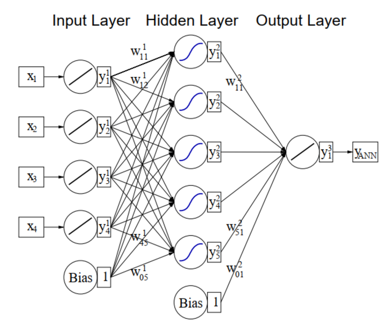


Deep neural networks (DNN)

single neuron

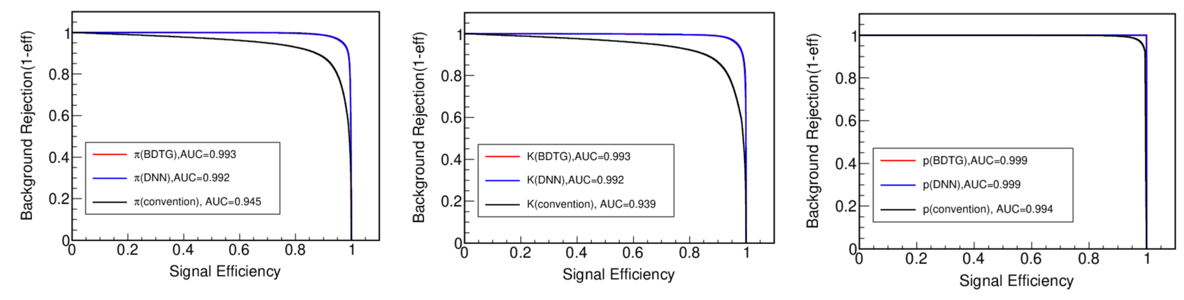


MultiLayer



- Receiver Operating Characteristic(ROC) curve

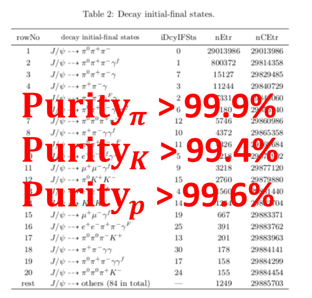
The performance of BDTG is slightly better than DNN!

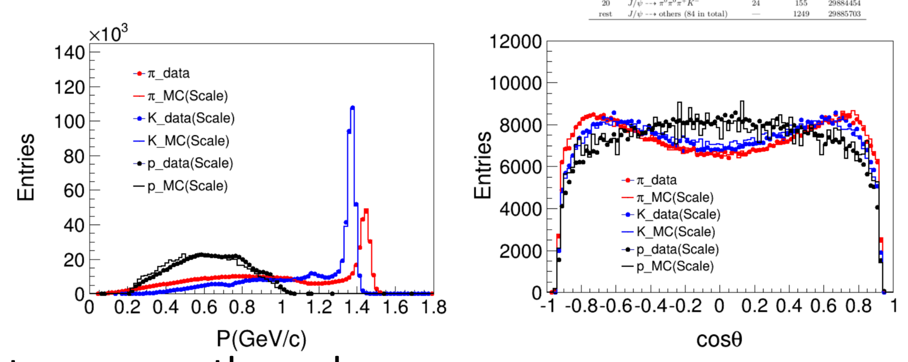


Dataset

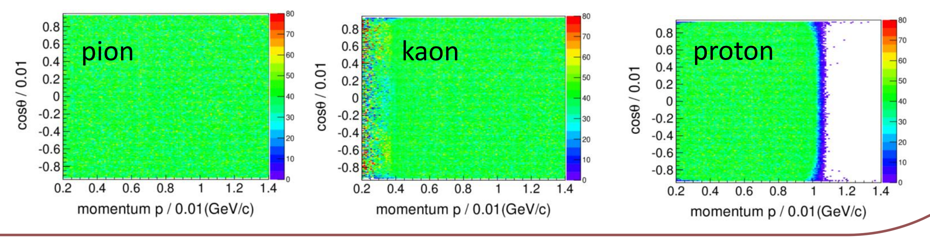
- The quality of data has a huge impact on machine learning tasks
- High statistic, wide range of momentum and solid angle and high purity hadron samples are required
- Event selection from 10 billion J/ψ events without conventional particle identification
  - * $J/\psi \to \pi^+\pi^-\pi^0, \pi^0 \to \gamma\gamma$
  - * $J/\psi \to K_S^0 K^\pm \pi^\mp, K_S^0 \to \pi^+\pi^-$
  - * $J/\psi \to p\bar{p}\pi^+\pi^-$

Topology analysis on MC samples





- It is essential to do data preprocessing, the steps we go through are:
  a) data cleaning
  b) data transformation
  c) class balance



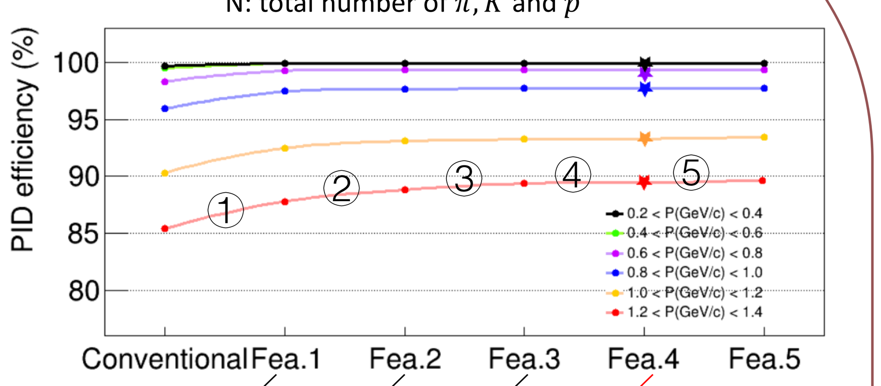
Feature selection

PID efficiency $= \frac{n}{N}$

n: number of π, K and p identified correctly

N: total number of π, K and p

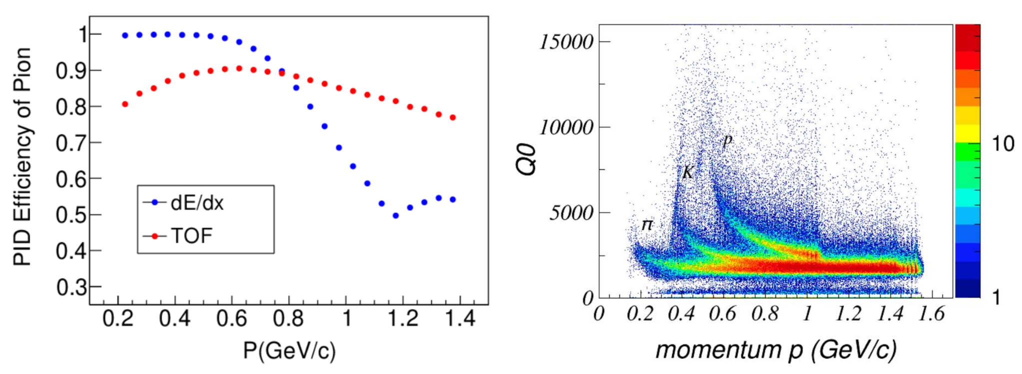
- It is stabilization for model training to get a subset of the most informative features from lots of available features:
- Boosted decision trees(BDTG) as select model
- Add features step by step based on detector understanding
- Select features according to importance, correlation matrix and particle identification efficiency



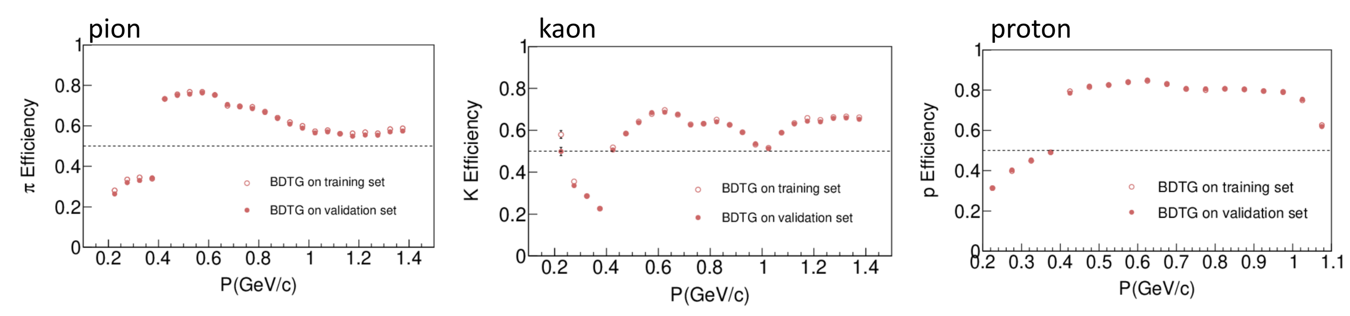
Fea.1	Fea.2	Fea.3	Fea.4	Fea.5
P	P	P	P	Fea.4
$cos\theta$	$cos\theta$	$cos\theta$	$cos\theta$	$nghits$
charge	charge	charge	charge	path
$\chi_{dE/dx}$	$\chi_{dE/dx}$	$\chi_{dE/dx}$	$\chi_{dE/dx}$	e3/e5
$t_{11,12,21,22}$	$t_{11,12,21,22}$	$t_{11,12,21,22}$	$t_{11,12,21,22}$	a42Mom
	Q_{TOF}	Q_{TOF}	Q_{TOF}	a20Mom
		E/P	E/P	$\Delta\varphi$
		eS/e3x3	eS/e3x3	Time
		secMom	secMom	dE
		latMom	latMom	energy
		$Nhits_{Emc}$	$Nhits_{Emc}$	Δx_{MUC}
		$\Delta\theta$	$\Delta\theta$	$\Delta\phi_{MUC}$
			depth	maxHit
				χ^2_{MUC}
				$Nhits_{MUC}$
				$NLay_{MUC}$

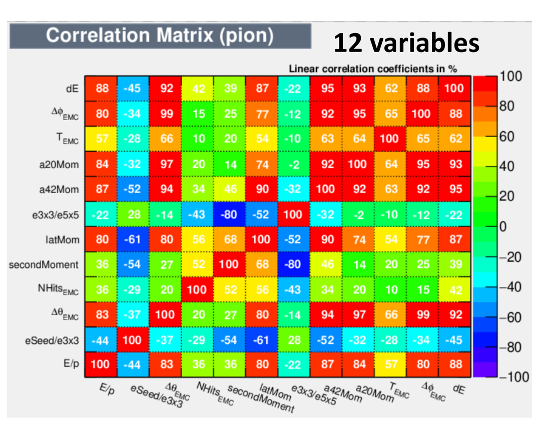
① The weights of dE/dx and TOF could be "learned"

② It is feasible to use Q_0 of TOF to identify the hadrons

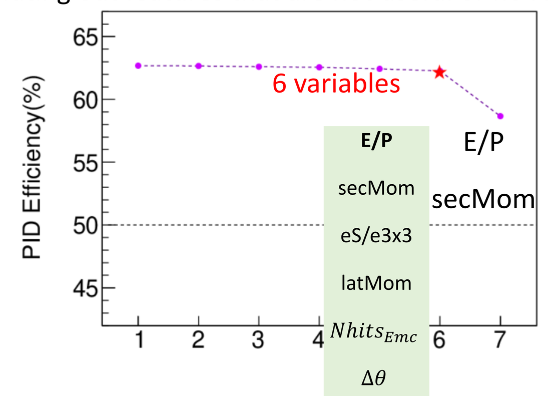


③ The measurements of EMC are also contribute to particle identification





Muting out features that are not useful or no relevant for PID

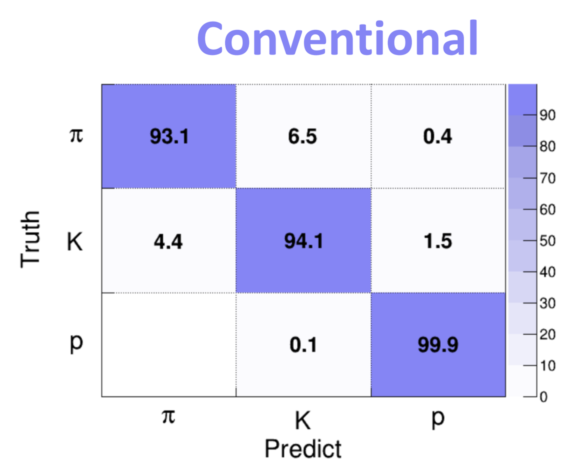


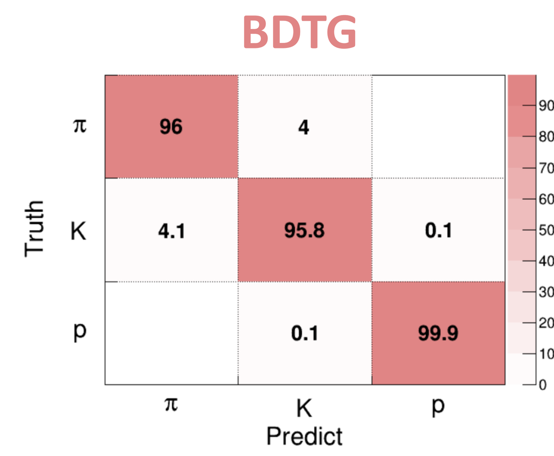
④ Depth measured by MUC are also be used for particle identification

⑤ Many other variables(Fea.5) are also tested, the result shows that more variables are useless for particle identification

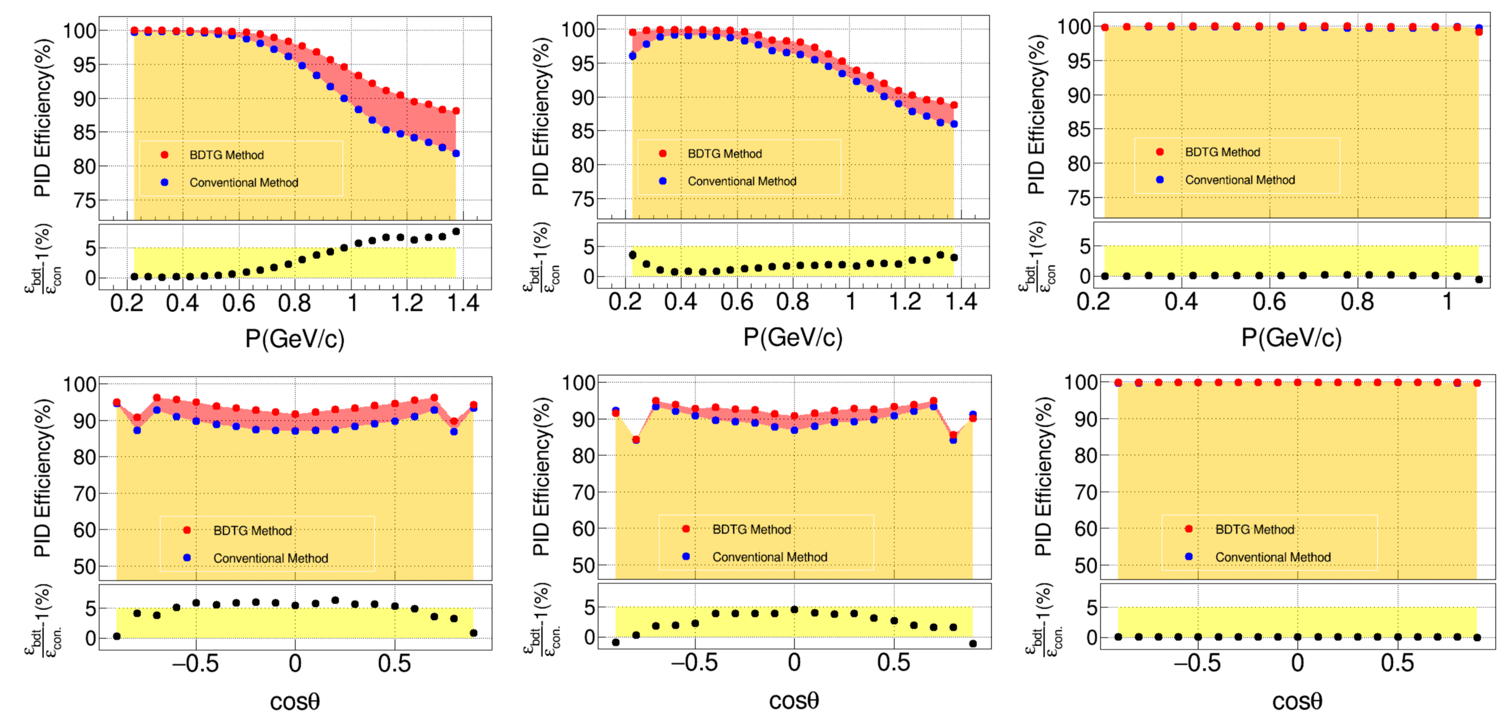
Performance

- Binary classification performed for particle identification with momentum between 0.2-1.4 GeV/c
- Cells show the ratio of particle labeled as column predicted as particle labeled as row
- Both for real data and MC samples, using BDTG method, true positive/negative rate improved and false positive/negative rate declined

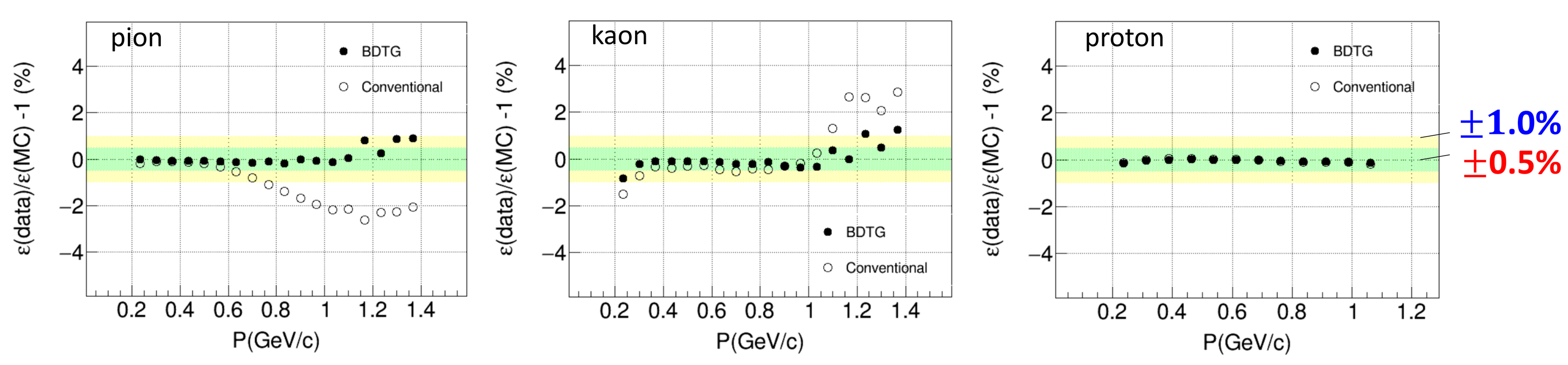




- The particle identification efficiency ε different with momentum and $\cos\theta$



- Systematic error: $\Delta\varepsilon = \frac{\varepsilon(data) - \varepsilon(MC)}{\varepsilon(MC)}$



- Systematic error of particle identification is improved using BDTG
  - * For the pion with momentum larger than 1.0GeV/c, systematic error decreased from 2% to 1%
  - * Below 1.1GeV/c, the systematic error less than 0.5% both for pion and kaon
  - * The systematic error of proton maintain very low

Conclusion

- With high quality data samples, particle identification are demonstrated using gradient boosted decision trees(BDTG) and deep neuron networks(DNN) method
- Combining information from the BESIII main drift chamber, time of flight detector, electromagnetic calorimeter and muon chamber allow to achieve higher particle identification efficiency for hadrons
- Training two models for real data and MC samples, systematic error of particle identification improved using boosted decision trees
- The application of more advanced machine learning techniques will be extended and optimized