Идентификации частиц с помощью машинного обучения на детекторе MPD

STudent Advanced Research Training at JINR

01.08.22 - 10.09.22 г. Дубна

Выполнил: Толкачёв Г.А.

07.10.2022

Введение

- Идентификация частиц является важным аспектом большинства экспериментов по физике элементарных частиц.
  - Идентификация долго живущих частиц, которые оставляют след в детекторе:
    электроны, мюоны, фотоны, заряженные пионы, заряженные каоны и тд.
  - Короткоживущие частицы идентифицируются по их распадам на долгоживущие частицы.

Цель работы: Выбор оптимальной модели MLP для улучшения эффективности идентификации заряженных частиц

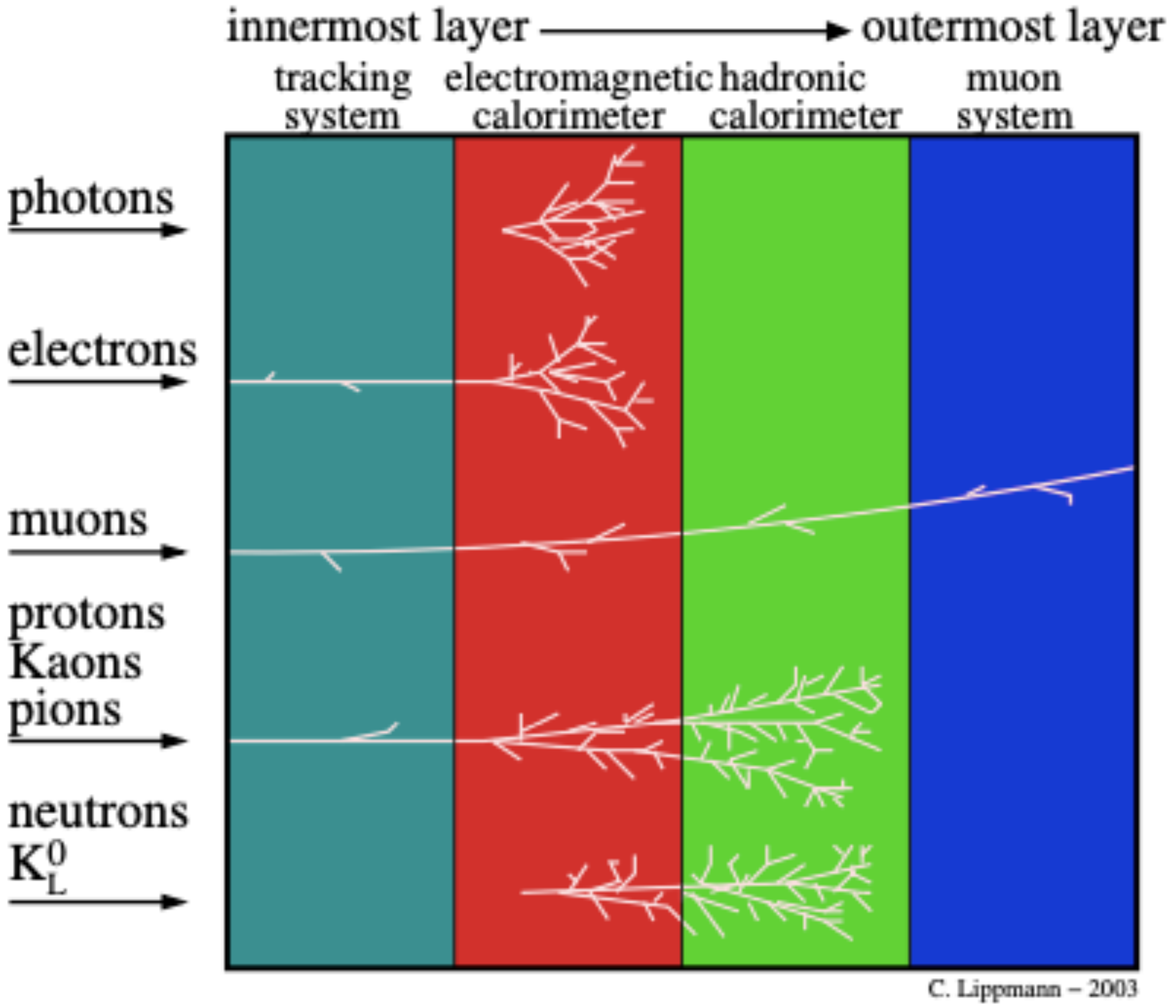
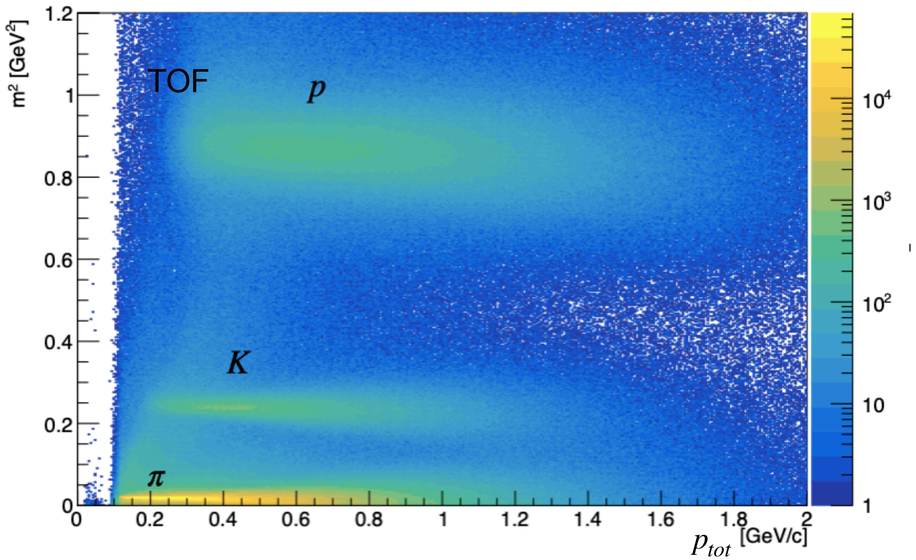


Схема детекторного комлекса

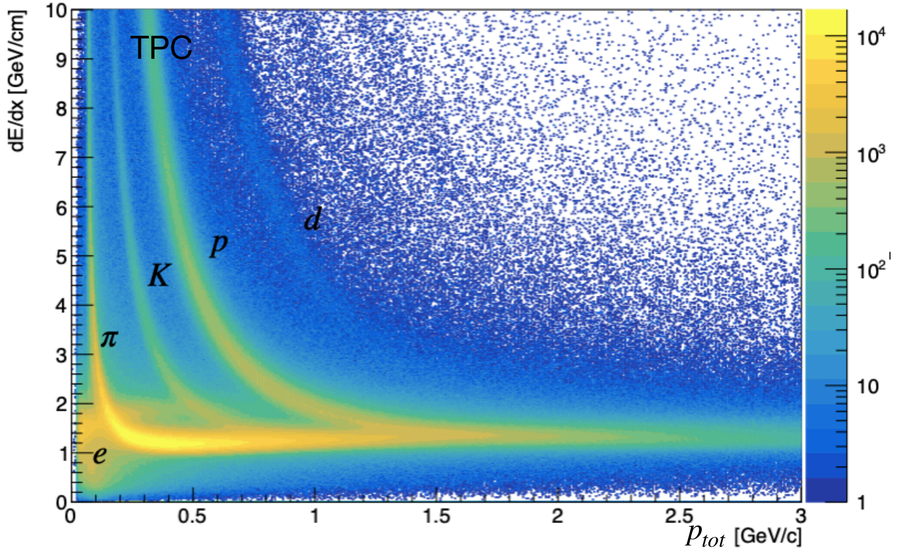
Отбор переменных

- Используемые переменные: dE/dx(dedx), m^2(m2), p_{tot}(momentum), η(eta), q(charge), nHints, dca, V_x, V_y, V_z.
- В исследовании было использовано 6 классов частиц: $\pi^-, \pi^+, K^-, K^+, p, \bar{p}$.
- Для обучения и тестирования моделей использовалось 200 000 событий для каждого класса

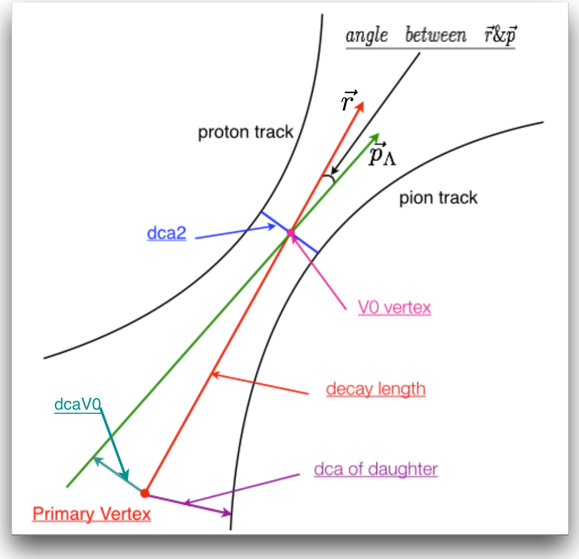
Распределение dE/dx как функция p_{tot} для π, K и протонов



Распределение dE/dx как функция p_{tot} для e, π, K, протонов и дейтронов



dca - расстояние ближайшего сближения



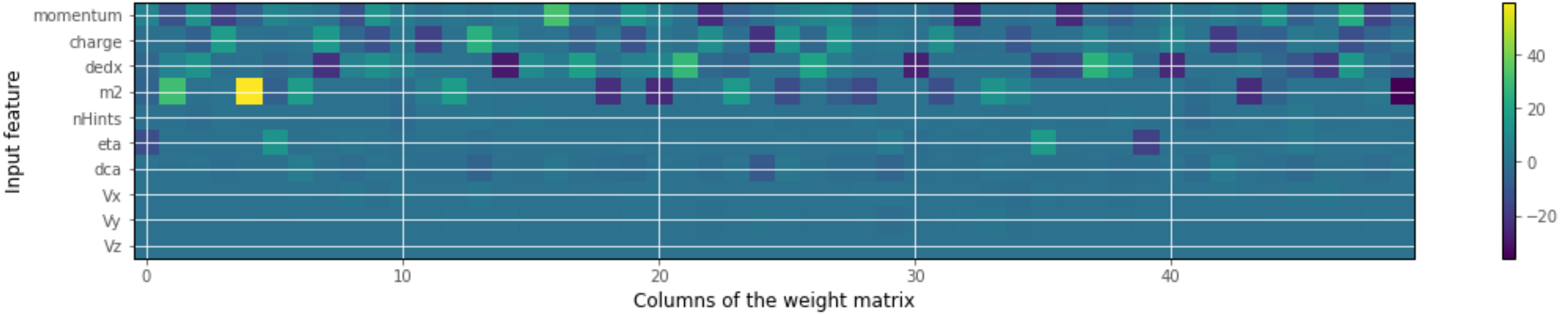
$$\beta = p/E = p/\sqrt{p^2 + m^2} \qquad (1)$$

$$m^2 = p^2(1/\beta^2 - 1) \qquad (2)$$

$$\frac{dE}{dx} = \frac{4\pi}{m_e c^2}\frac{nz^2}{\beta^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\left[ln\left(\frac{2m_e c^2\beta^2}{I(1-\beta^2)}\right) - \beta^2\right] \qquad (3)$$

Отбор переменных

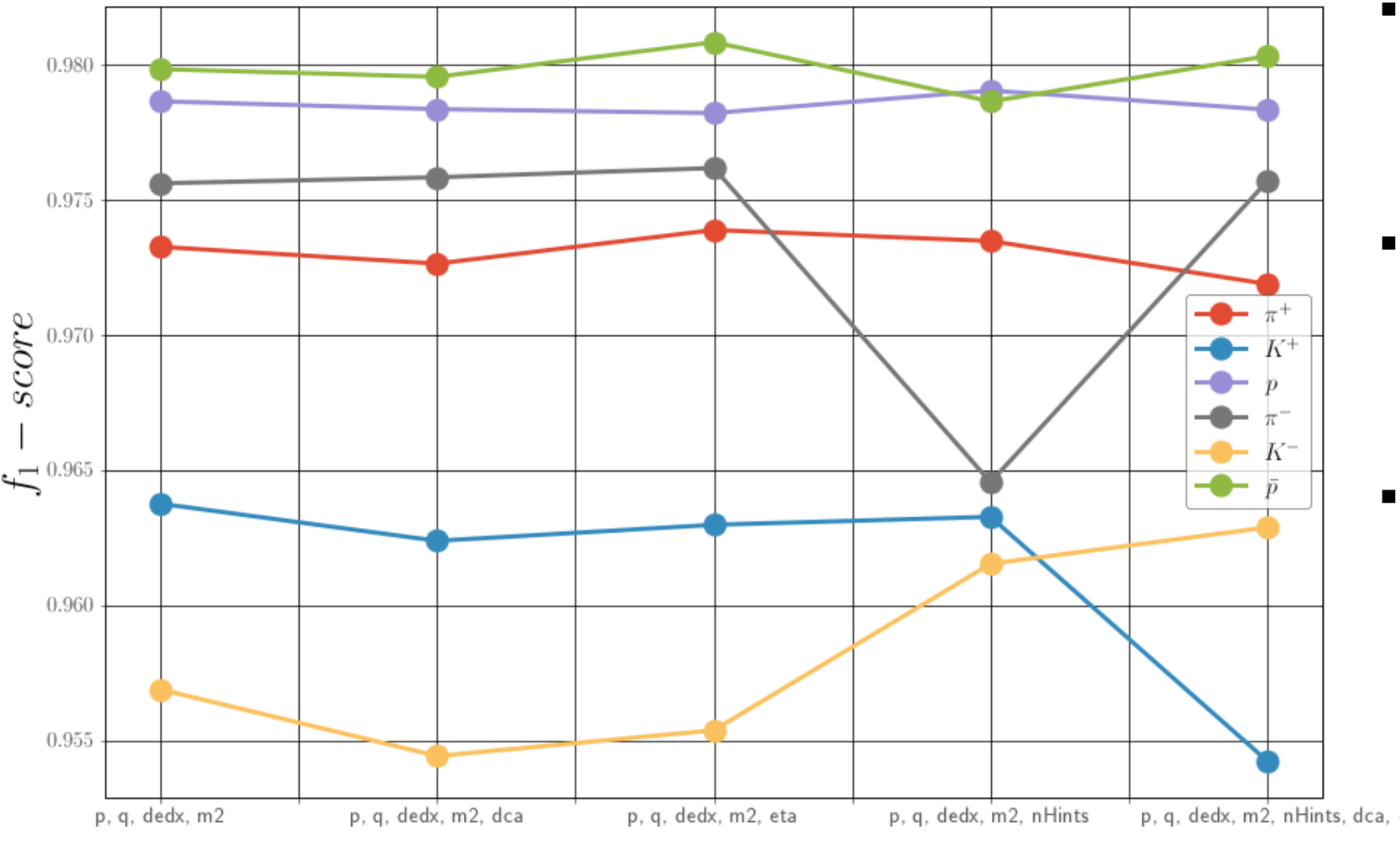
Матрица весов



- Переменные: dedx, m2, momentum, charge почти для кажого элемента скрытого слоя имеют вес отличный от нуля
- Переменные: nHints, dca, eta для некоторых элементов скрытого слоя имеют вес отличный от нуля
- Vx,Vy,Vz - имеют нулевой вес для всех элементов скрытого слоя

Отбор переменных

Зависимость f1-score от набора переменных



- Причина, по которой $K^{\pm}$ имеют самый низкий показатель f1-score, заключается в том, что, например, на распределении m^2 они находится между p и π и смешивается со всеми из них
- Некоторые дополнительные переменные улучшают f1-score для одного типа частиц и ухудшают для другого типа. Остальные дополнительные переменные не вносят значительного вклада в f1-score
- Далее в работе использовался набор параметров: momentum, charge,dedx, m2

$$f_1 = 2 * \frac{recall * precision}{recall + precision} \quad (4)$$

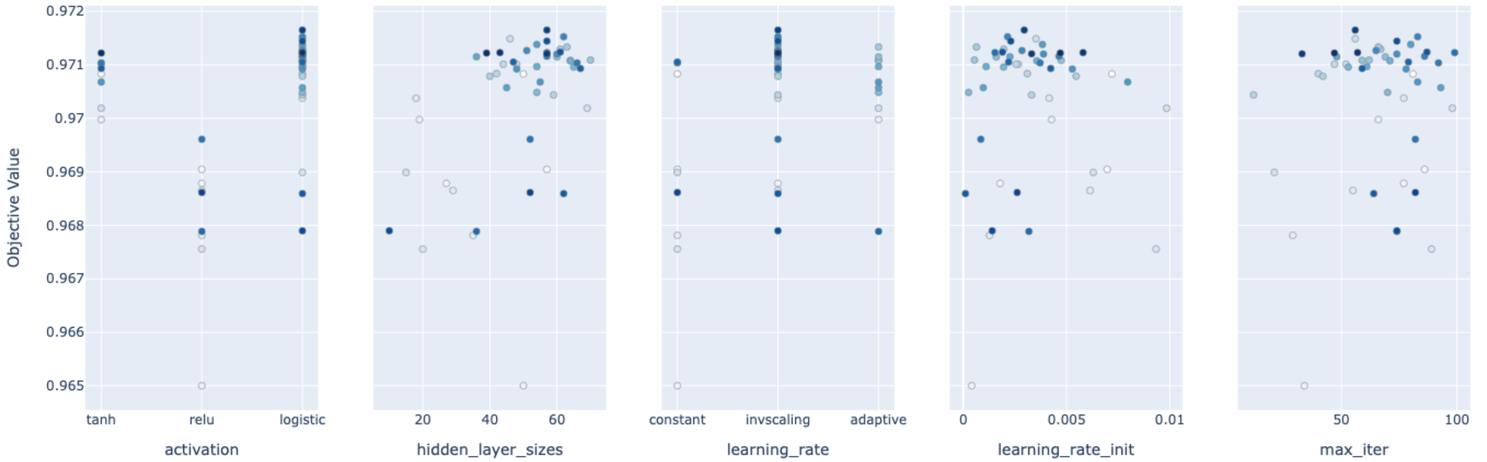
$$precision = \frac{TP}{TP + FP}, \quad (5) \quad recall = \frac{TP}{TP + FN} \quad (6)$$

Оптимизация гиперпараметров

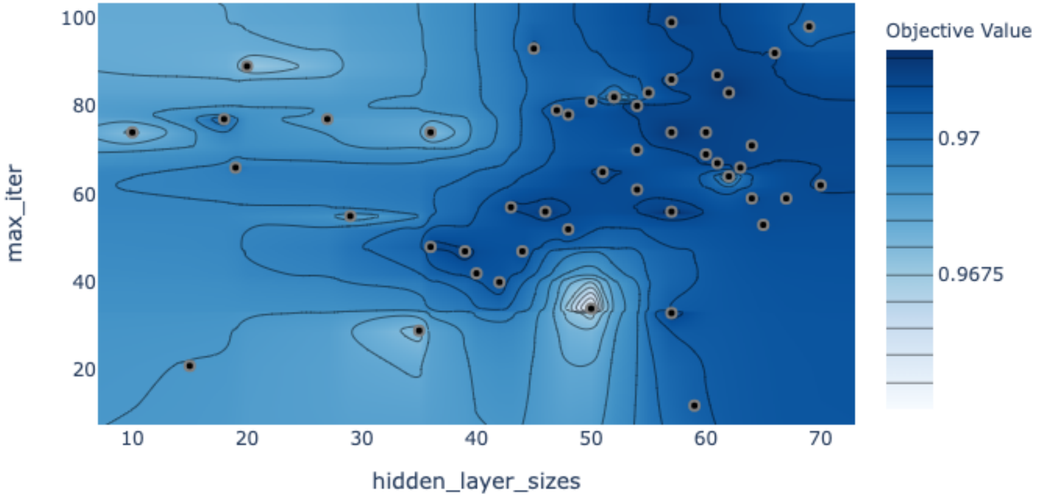
Набор гиперпараметров, которые использовались в байесовской оптимизации

hidden_layer_sizes	10 - 70
max_iter	10 - 100
learning_rate_init	0.0001 - 0.01
activation	logistic, tanh, relu
learning_rate	constant, invscaling, adaptive

Зависимость f1-score от значения гиперпараметра



Карта гиперпараметров



Оптимальный набор параметров

hidden_layer_sizes	36
max_iter	48
learning_rate_init	0.006
activation	logistic
learning_rate	constant

- Больше чиссло классификаторов имеют f1-score > 0.97
- Для упрощение модели и снижения вычислительныз затрат выбрана модель с hidden_layer_sizes = 36 и max_iter = 48
- Может быть выбрано меньшее число max_iter

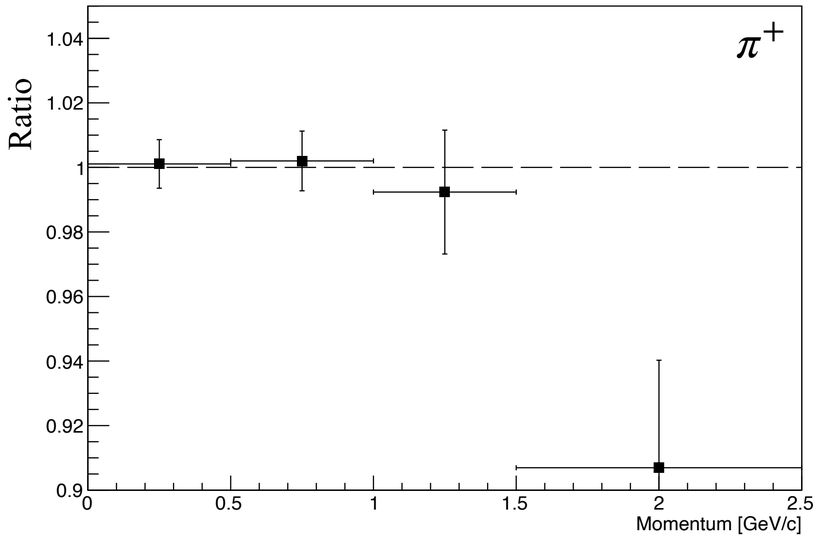
Дополнительные исследования

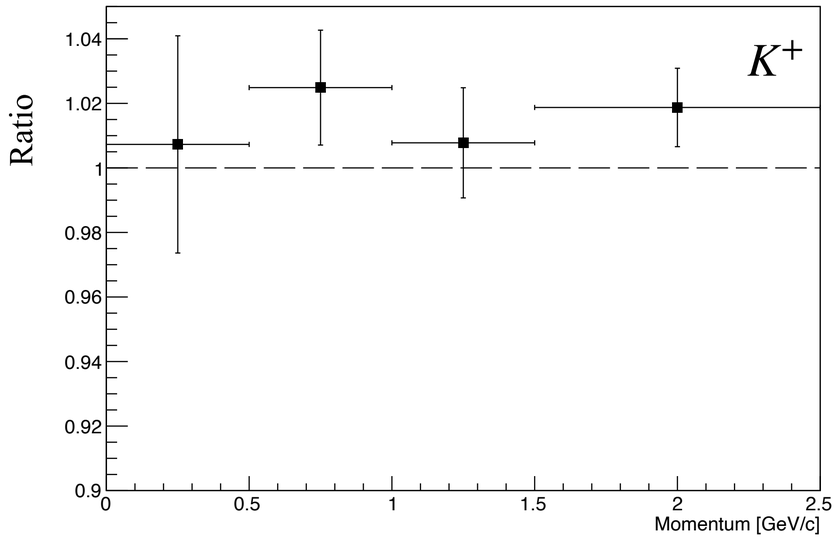
Обучение в разном диапазоне импульса

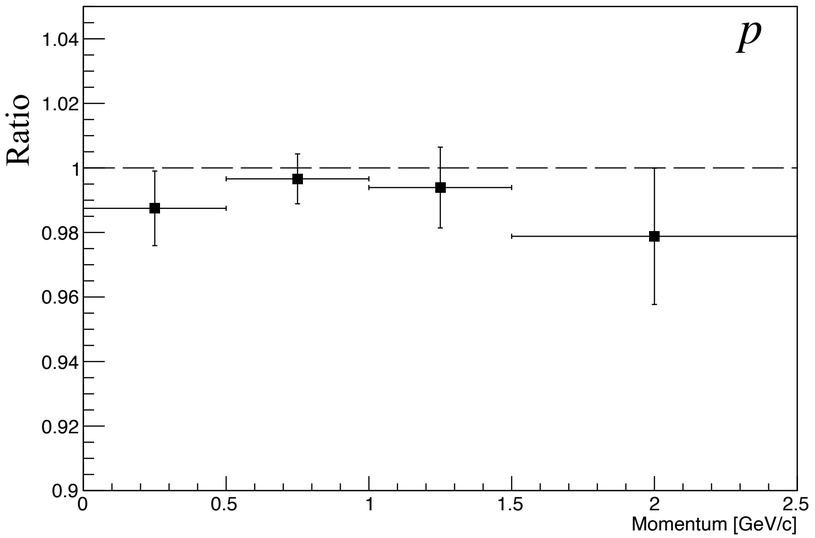
- Идея: обучить несколько моделей MLP с использованием данных из разных диапазонов импульса.
- Используемые диапозоны импульсов: [0.0, 0.5, 1.0, 1.5, 2.5]
- Оценка качества данного подхода производилась с помощью отношения верных ответов от набора моделей обученных в разных диапазонах импульса и от модели, которая была обучена на всем интервале импульса:

$$Ratio = \frac{N_{true}^{i-\text{range model}}}{N_{true}^{\text{full range model}}} \quad (7)$$

- Данный подход не вносит значительного вклада в идентификацию частиц.







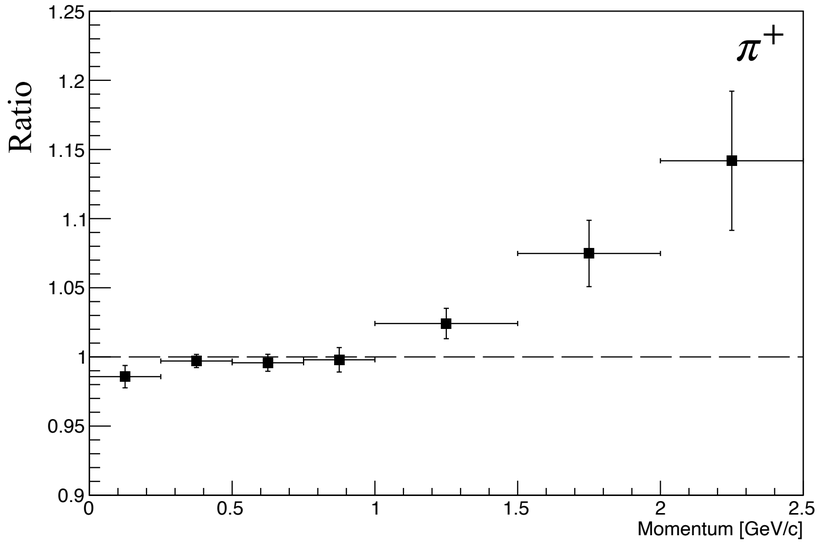
Дополнительные исследования

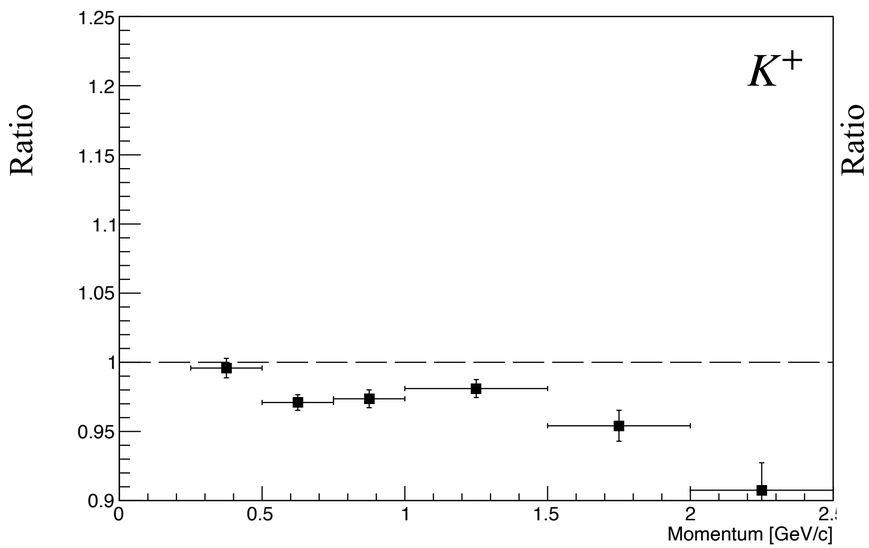
Бинарная классификация для каждого класса

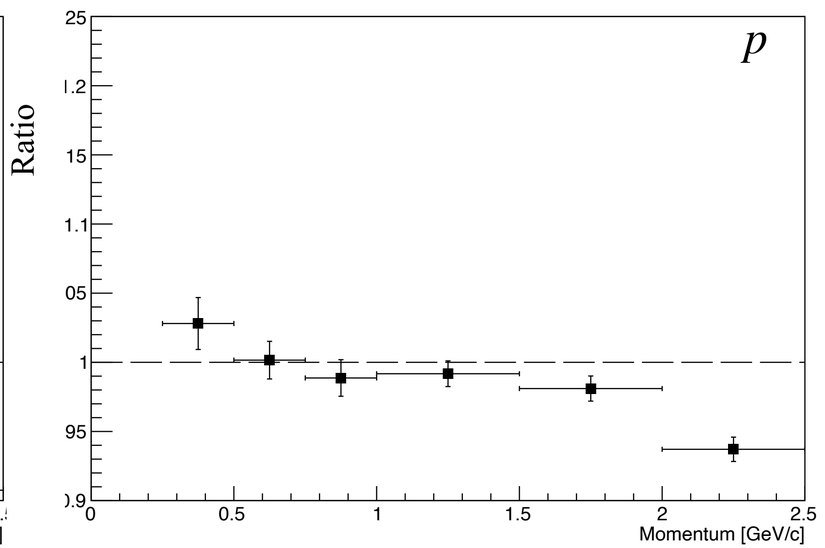
- Идея: для каждого класса частиц обучить бинарную модель MLP.
- Оценка качества данного подхода производилась с помощью оценки отношения верных ответов от набора бинарных классификаторов и от мульти классификатора:

$$Ratio = \frac{N_{true}^{\text{binary classification}}}{N_{true}^{\text{multiclass classification}}} \quad (8)$$

- Данный подход не вносит значительного вклада в идентификацию частиц.







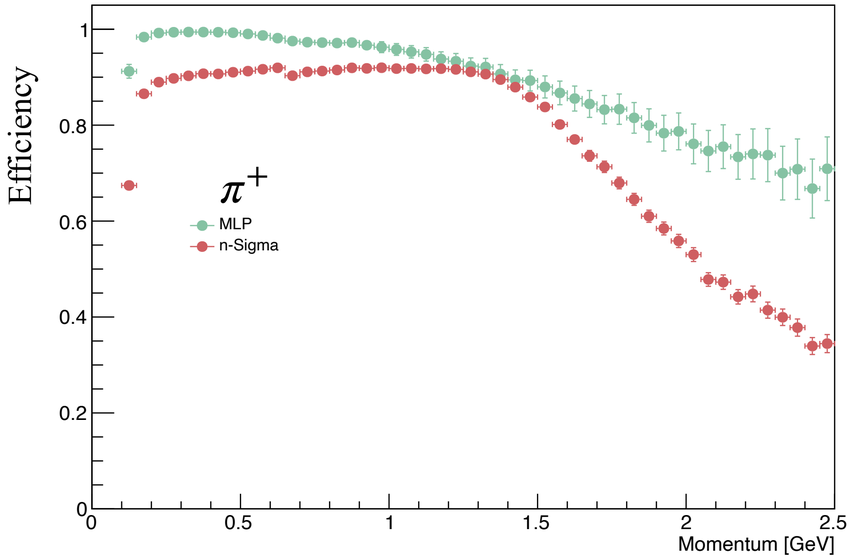
Результаты

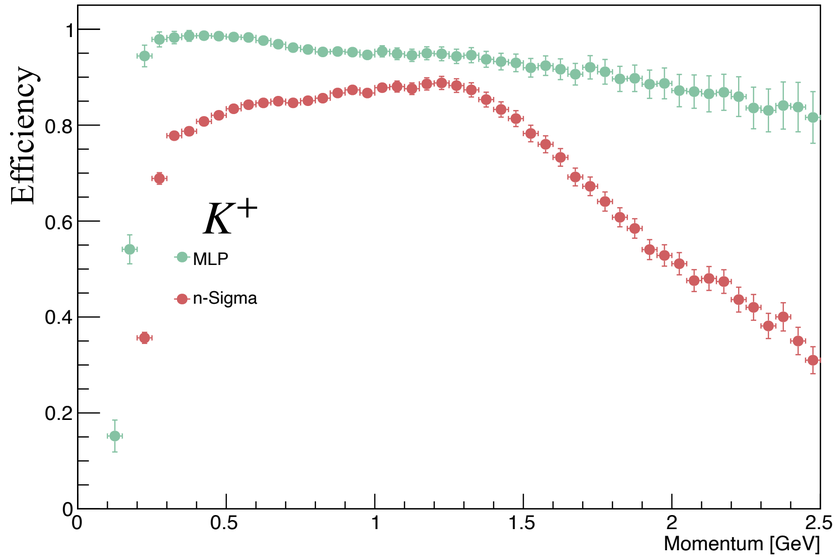
- Для оценки качетсва идентификации использовалась эффектиность:

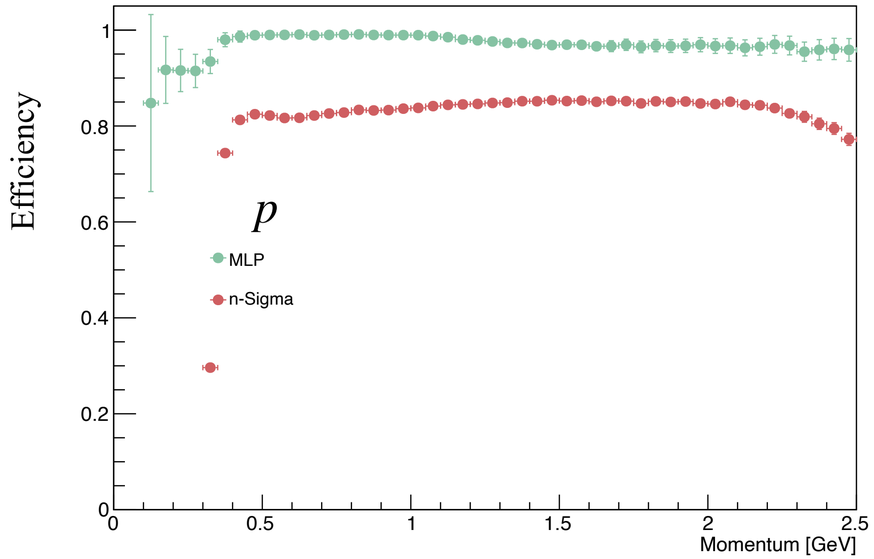
$$Efficiency = \frac{dN^i_{\text{true}}/dp}{dN^i_{\text{all gen.}}/dp} \qquad (9)$$

- Эффективность идентификации модели MLP сравнивается с эффективностью идентификации стандартного n-Sigma подхода

- Для каждого класса частиц MLP подход имеет эффективность идентификации выше чем n-Sigma подход во всем интервале импульса







Заключение

- Для классификатора MLP было выбрано количество переменных, которые вносят наибольший вклад в идентификацию.

- Используя байесовскую оптимизацию были выбраны гиперпараметры, которые не усложняют MLP-модель и позволяют получить высокий f1-score.

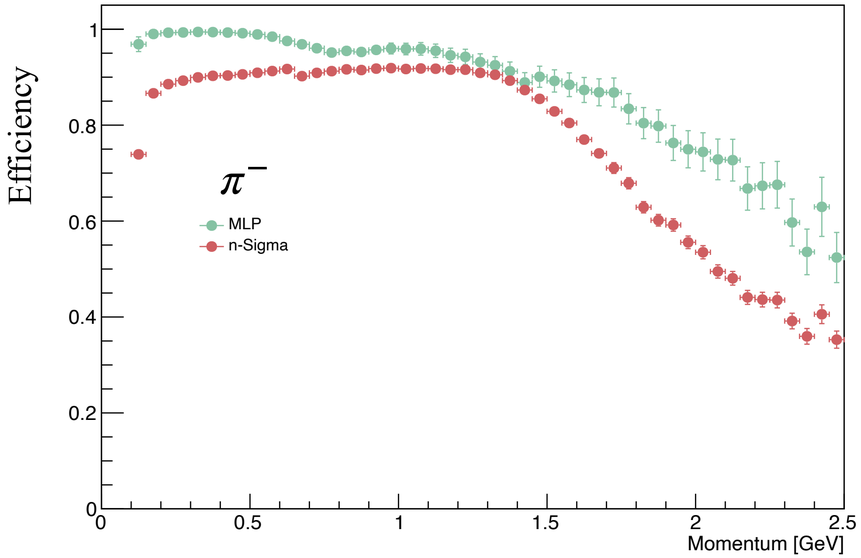
- Исследованы дополнительные подходы для улучшения качества правильной классификации частиц. Каждый из них не показал существенных результатов, однако эти подходы могут быть исследованы в будущем с другой настройкой.

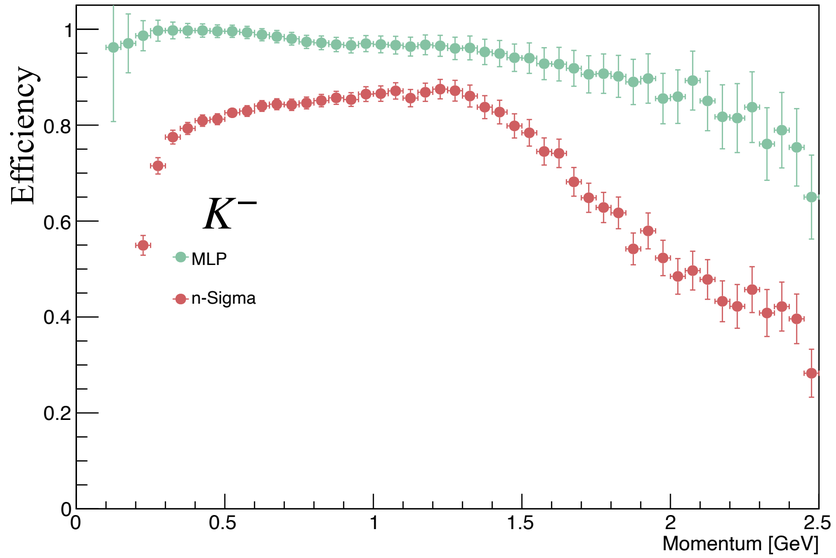
- Был изучен n-sigma подход и проведено сравнение с MLP-подходом для идентификации частиц. Было показано, что использование классификатора MLP для идентификации частиц значительно повышает эффективность для каждого вида частиц.

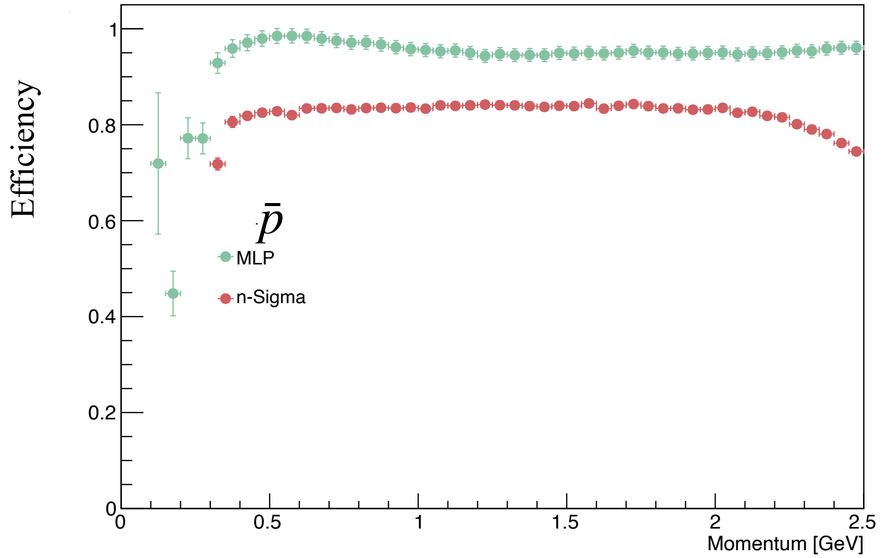
Спасибо за внимание!

Дополнительные слайды

Результаты


Efficiency
π^-
MLP
n-Sigma
Momentum [GeV]

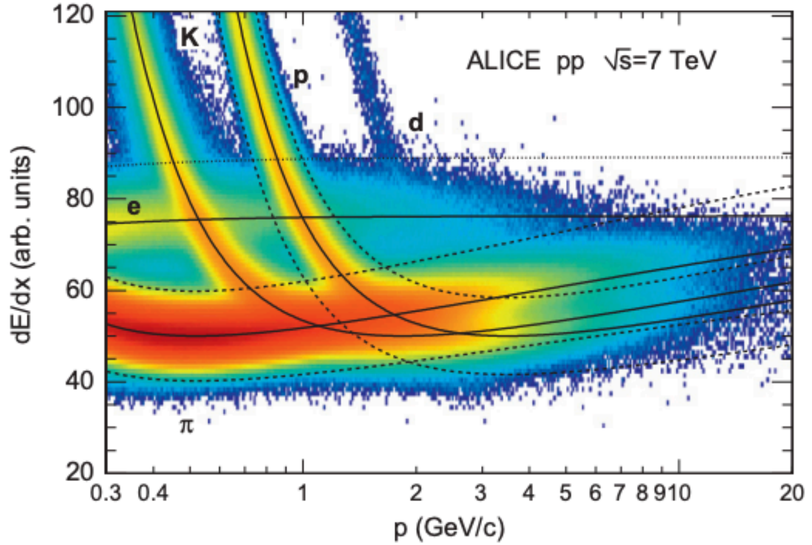

Efficiency
K^-
MLP
n-Sigma
Momentum [GeV]


Efficiency
$\bar{p}$
MLP
n-Sigma
Momentum [GeV]

Медоты идентификации

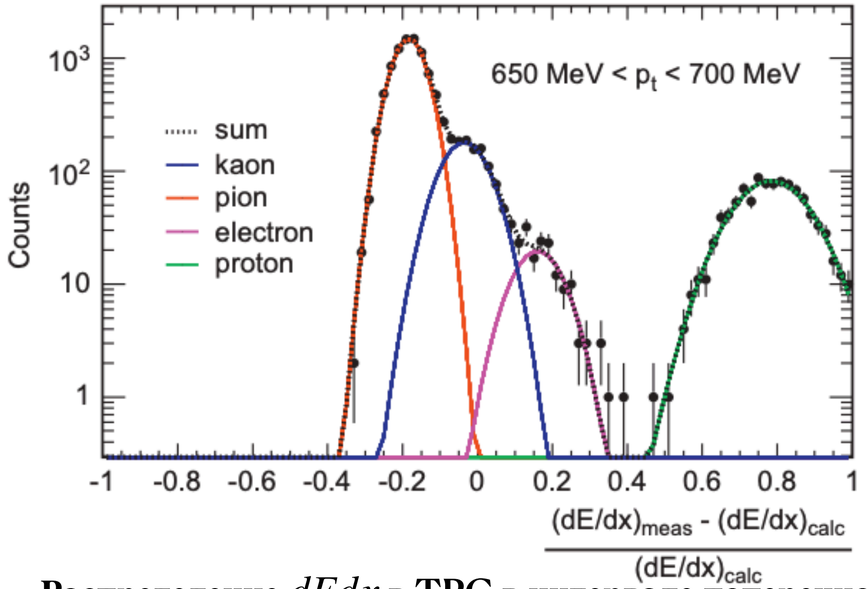
Параметрицизия формулы Бете-Блоха

- Для применения метода идетификации необходимо идентификации параметризацию формулы Бете-Блоха, которая связывает ионизационные потери энергии заряженной частицы с ее скоростью.
  - $$\frac{dE}{dx}(\beta\gamma) = \frac{P_1}{\beta^{P_4}}\left(P_2 - \beta^{P_4} - ln\left(P_3 + \frac{1}{(\beta\gamma)^{P_4}}\right)\right) \quad (10)$$
- Далее необходимо произвести оценку параметров $P_1 - P_5$ для каждого семейтсва частиц.



Распределение $dEdx$ в зависимости от импульса трека детектороном комлексе ALICE TPC[1].

[1] https://doi.org/10.1016/j.nima.2012.05.022



Распределение $dEdx$ в TPC в интервале поперечного импульса трека $650MeV/c < p_T < 700MeV/c$.

Медоты идентификации

$n\sigma$ подход

- Наиболее часто используемой различающей переменной для идентификации частиц является переменная N_{σ^i}, определяемая как отклонение измеренного сигнала от наиболее вероятного значения для каждого вида частиц i. Для TPC и TOF N_{σ}^{i} определяется как:

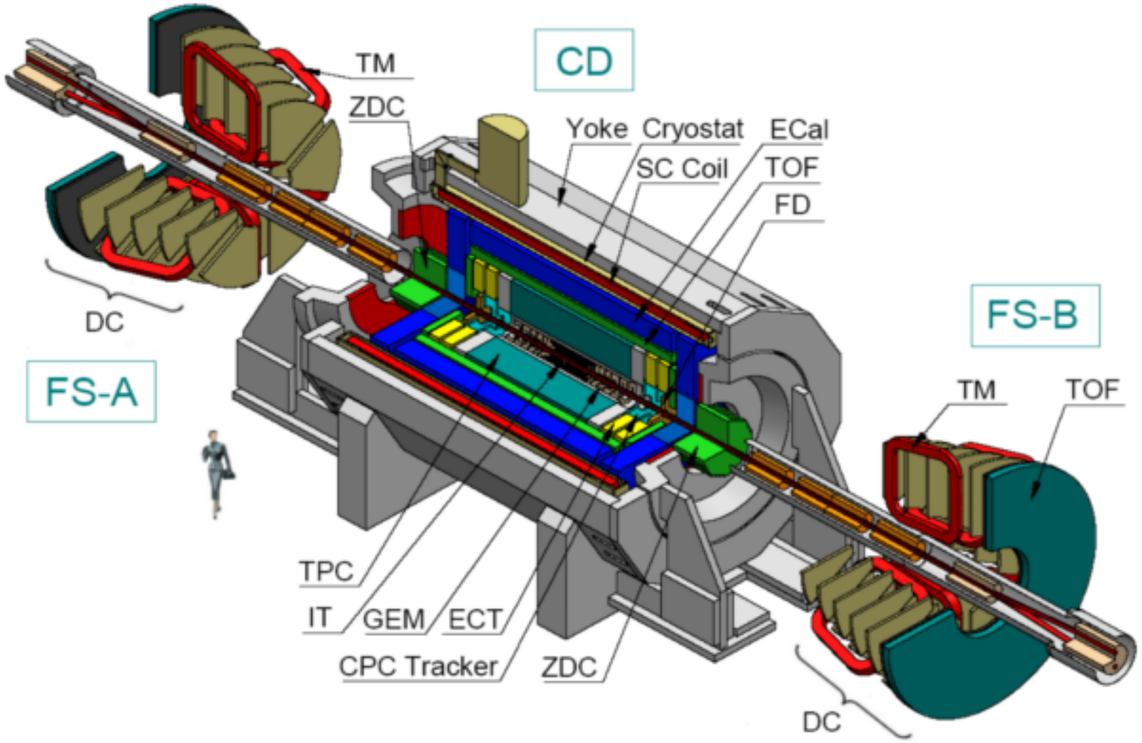
$$N_{\sigma_{TPC}}^{i} = \frac{dE/dx - \langle dE/dx \rangle^{i}}{\sigma_{TPC}^{i}}, \quad (11) \qquad N_{\sigma_{TOF}^{i}} = \frac{m^2 - \langle m^2 \rangle^{i}}{\sigma_{m^2}^{i}}, \quad (12)$$

- Частица идентифицируется как частица определенного типа, если это значение находится в определенном диапазоне вокруг математического ожидания $N_{\sigma_{TPC}} = 2$и $N_{\sigma_{TOF}} = 2$ (Можно выбрать другое значение).

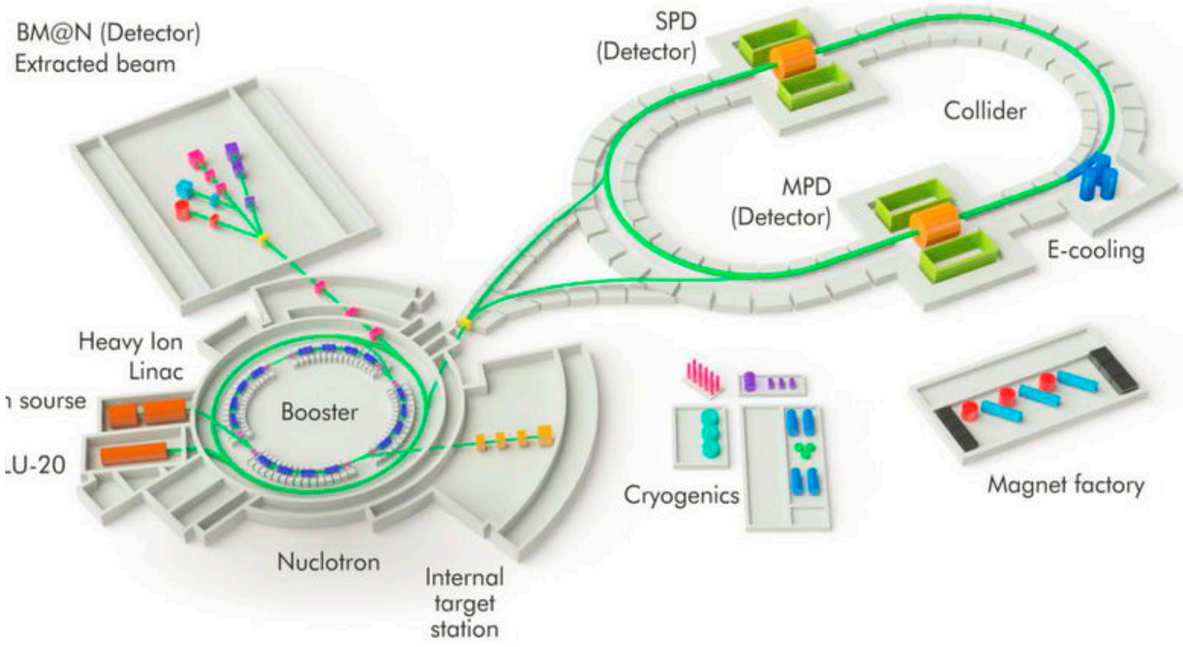
$$N_{\sigma} \leq \sqrt{N_{\sigma_{TOF}^{i}}^{2} + N_{\sigma_{TPC}^{i}}^{2}}, \quad (13)$$

- Если условие (13) выполнено для N_{TPC}^{i} и N_{TOF}^{i}, то частица идентифицируется как i-вид. В случае, если частица может быть совместима с более чем одним видом, подход $n\sigma$ соответствует ложному решению.

MPD


TM
ZDC
CD
Yoke
Cryostat
ECal
SC Coil
TOF
FD
DC
FS-A
FS-B
TM
TOF
TPC
IT
GEM
ECT
CPC Tracker
ZDC
DC

NICA


BM@N (Detector)
Extracted beam
SPD
(Detector)
Collider
MPD
(Detector)
E-cooling
Heavy Ion
Linac
Booster
Nuclotron
Internal
target
station
Cryogenics
Magnet factory

Использованные данные

Смоделированные данные, использованные в работе, были получены методом Монте-Карло с использованием генераторов UrQMDv3.4 при условии реальных столкновений Bi-Bi эксперимента MPD с $\sqrt{s_{NN}}$=9,2 ГэВ.

Критерии на отбор событий.

p_{tot}	$\|\eta\|$	r	nHints	dca	|Vz|
>0.1 GeV	<1.5	<1.25 cm	> 15	< 5 cm	< 100 cm