

Яндекс



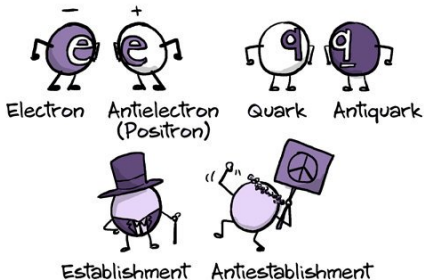
Идентификация частиц в детекторе LHCb в ЦЕРНе

Никита Казеев^{1,2}

¹Национальный исследовательский университет "Высшая школа экономики" ²Яндекс

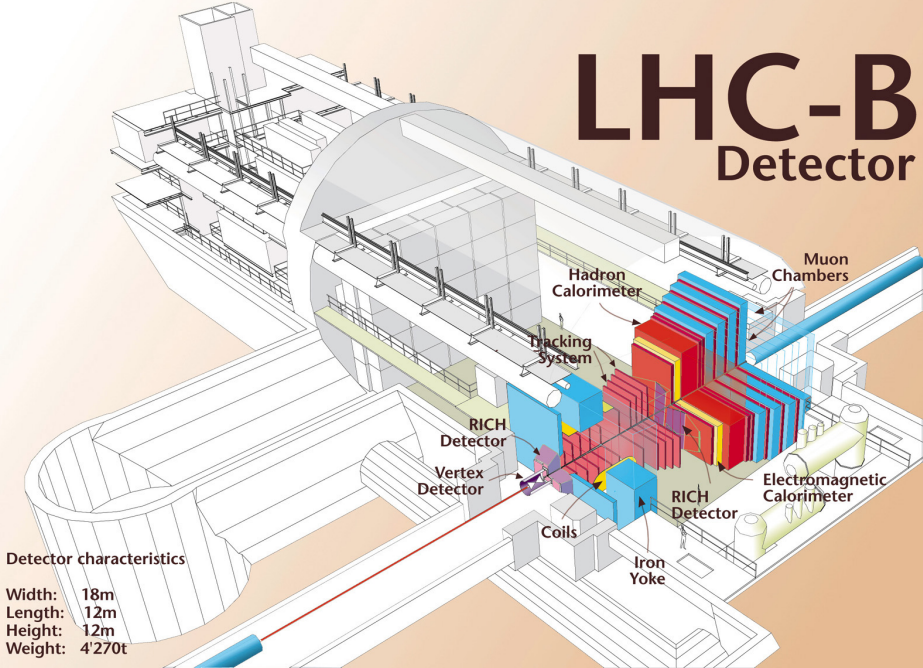
Дисбаланс материи и антиматерии

Particles are Political



"We Have No Idea"
Book available now!
wehavenoidea.com

LHC-B Detector

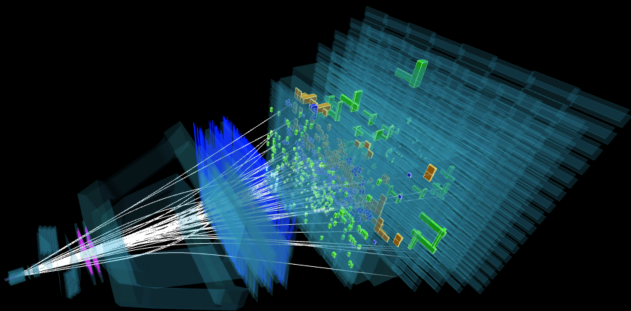


Идентификация частиц



NO PID

Event 2717213713
Run 199643
Thu, 28 Sep 2017 10:39:17



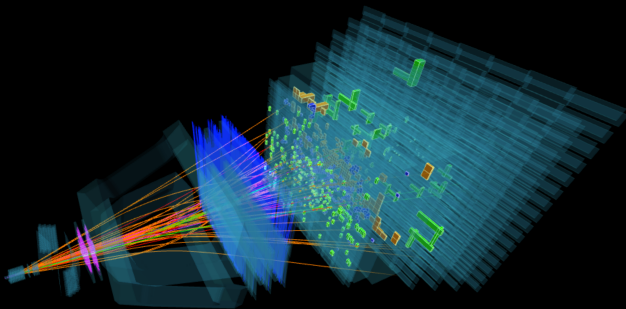


Open Controls

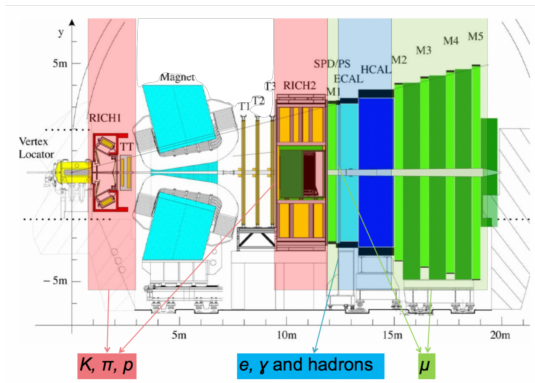
Event 2717213713

Run 199643

Thu, 28 Sep 2017 10:39:17

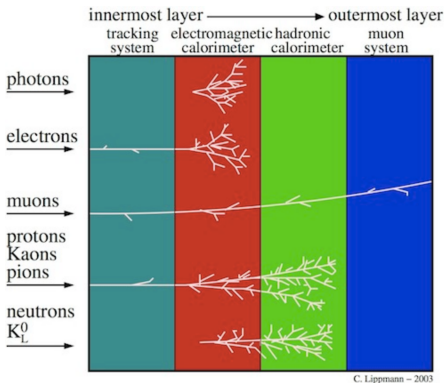


Аппаратные средства



- › Трекер
- › Черенковский детектор
- › Калориметр
- › Мюонные камеры

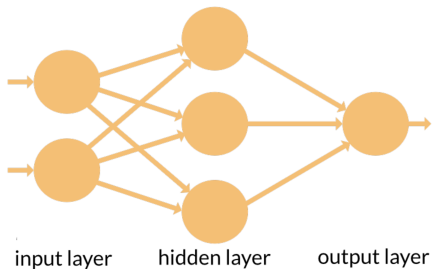
Задача



- › Типы частиц: электрон, мюон, пион, каон, протон, “приведение” (ошибка реконструкции)
- › Высокоуровневые признаки: выходы экспертных алгоритмов по подсистемам
- › Данные: 12 миллионов симулированных событий

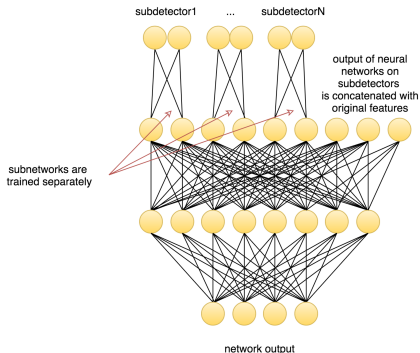
Baseline

- › Нейросеть с одним скрытым слоем
- › 6 моделей 1-vs-all



Пробовали: хитрые нейросети

- › Отдельные “головы”, обучающиеся на отдельных субдетекторах
- › Общая нейросеть, использующая их выходы плюс сырые признаки



Лучшее решение

- › Градиентный бустинг (CatBoost)
- › (На самом деле kaggle-style ensemble всех моделей вместе, но прирост качества не стоит роста вычислительных затрат)



Результаты. AUC one-vs-rest

Particle type	Ghost	Electron	Muon	Pion	Kaon	Proton
Baseline	0.9484	0.9855	0.9844	0.9346	0.9148	0.9178
XGBoost	0.9609	0.9908	0.9922	0.9568	0.9303	0.9302
Decision Train	0.9637	0.9914	0.9927	0.9577	0.9309	0.9310
CatBoost	0.9664	0.9917	0.9929	0.9586	0.9322	0.9323
One-layer NN	0.9615	0.9910	0.9922	0.9574	0.9305	0.9304
Deep NN	0.9632	0.9915	0.9925	0.9587	0.9320	0.9319
Stack NN	0.9623	0.9911	0.9924	0.9578	0.9315	0.9312
Special NN	0.9621	0.9910	0.9923	0.9576	0.9308	0.9301
Flat 2d grid	0.9594	0.9868	0.9865	0.9494	0.8978	0.9096
Flat 2d	0.9605	0.9883	0.9887	0.9514	0.9146	0.9139
Flat 4d	0.9600	0.9875	0.9877	0.9505	0.9130	0.9130

Продолжение следует

- › Предложенные методы позволяют значительно улучшить качество идентификации частиц

Продолжение следует

- › Предложенные методы позволяют значительно улучшить качество идентификации частиц
- › **НО.** Модернизация LHCb увеличит число треков на порядок

Продолжение следует

- › Предложенные методы позволяют значительно улучшить качество идентификации частиц
- › **НО.** Модернизация LHCb увеличит число треков на порядок
- › Соответственно возрастёт сложность задачи идентификации частиц

Благодарности

- › Андрею Устюжанину за научное руководство
- › Михаилу Гущину и Денису Деркачу, вместе с которыми мы делаем этот проект