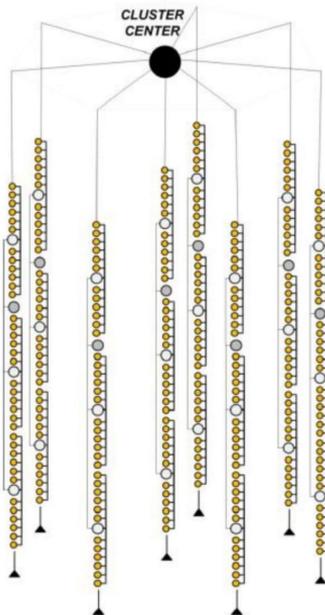
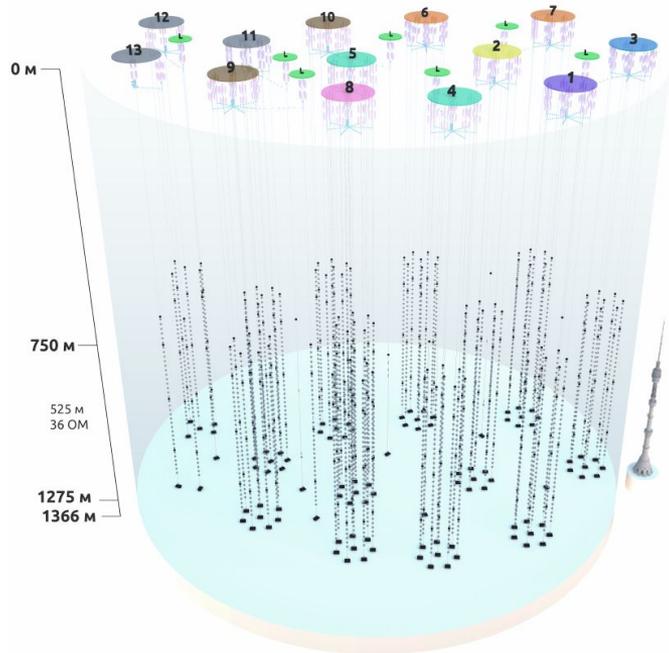


Машинное обучение в Байкал-GVD



Байкал-GVD



Детектор:

эффективный объем: 0.6 км^3

14 кластеров:

почти правильный семиугольник,
радиус: 60 м,
межкласт. расстояние: 300 м

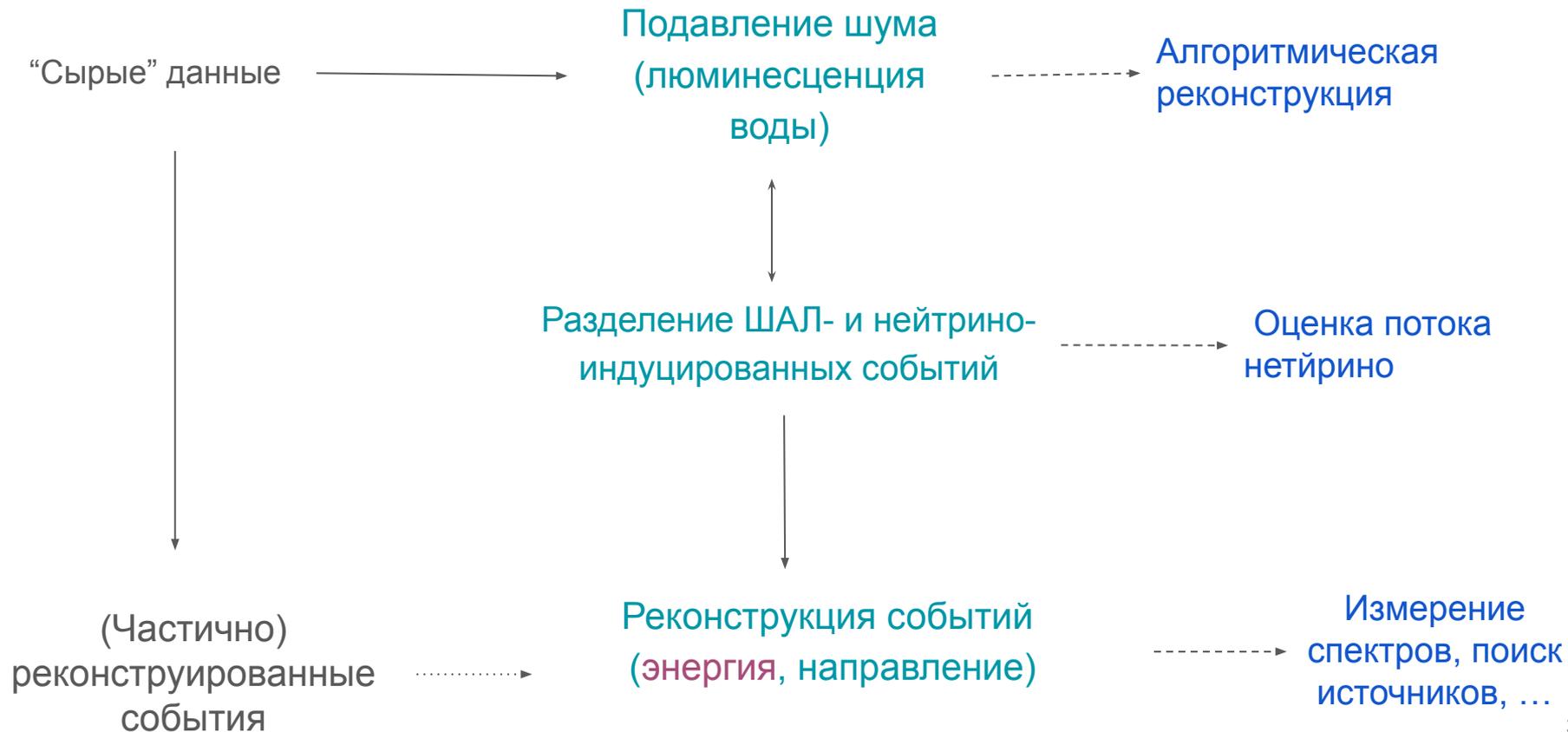
8 гирлянд (на кластер):

рабочая часть: от 750 до 1275 м,
разделены на несколько секций

**36 оптических модулей
(на гирлянде):**

содержат по 1 ФЭУ,
смотрят вниз,
шаг установки ОМ: 15 м,
Точность определения:
положений ОМ: $\sim 20 \text{ см}$,
времен срабатывания ОМ: $\sim 2 \text{ нс}$.

Общий план применения МО



Используемое Монте-Карло

Два типа событий:

- 1) Мюоны от космических лучей (μ) ([сверху](#))
- 2) ν_μ -индуцированные события ([снизу](#))

Полная симуляция развития ШАЛ,
параметрическая эволюция ливней в воде.

Спектр МК событий совпадает с
ожидаемым экспериментально

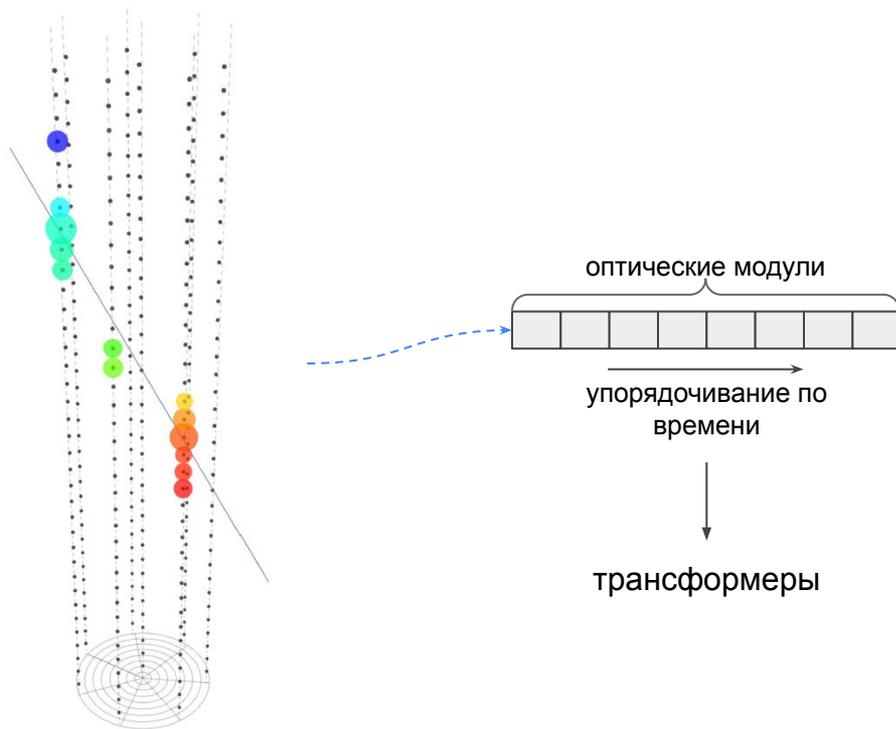
Фиксированная геометрия (небольшие
флуктуации)

Временная развертка сигнала заменяется
на дискретные хиты

Для события известны: параметры первичной частицы (энергия, угол).

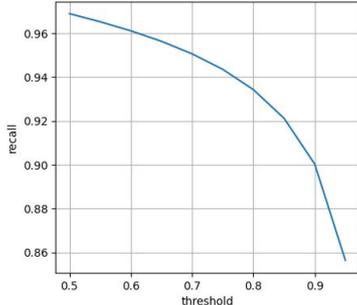
Для хитов известны: координаты x, y, z ; время активации; интегральный сигнал;

Представление данных для нейронных сетей

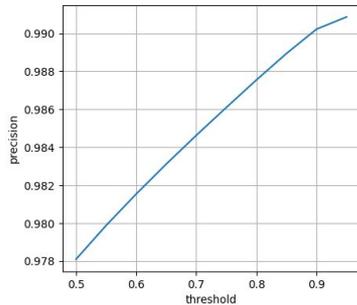


Подавление шумовых срабатываний ОМ

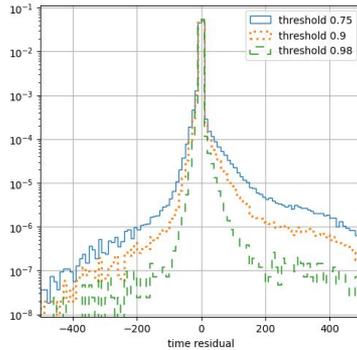
Точность



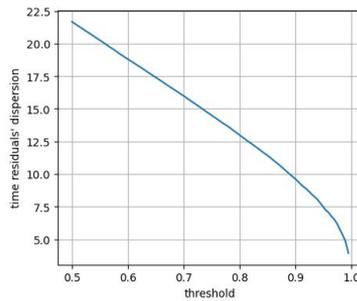
Полнота



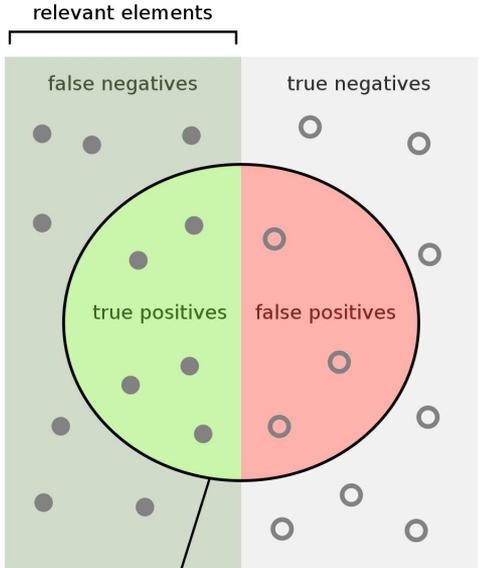
Распределение t_{res}



t_{res}



Скорость обработки быстрее до 10^3+ раз.



How many retrieved items are relevant?

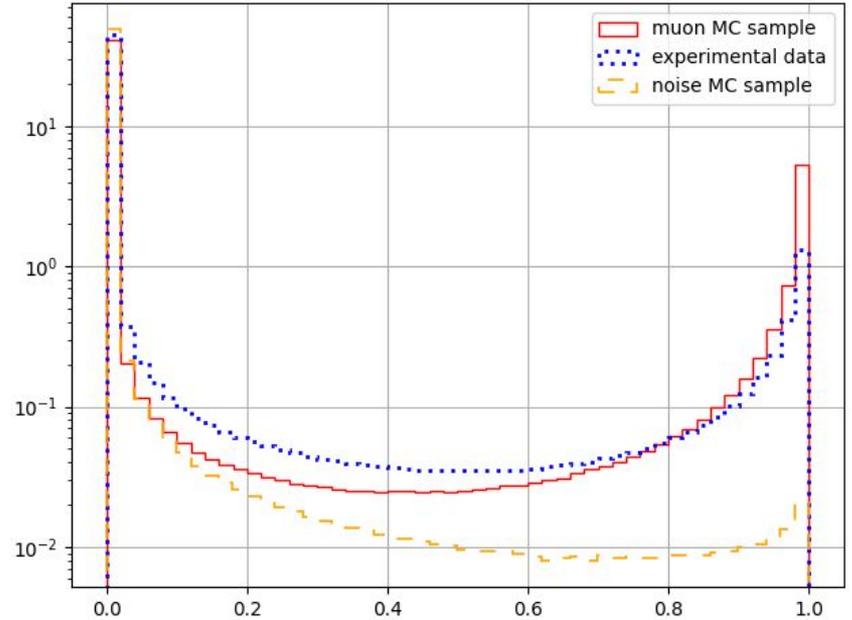
$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are retrieved?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

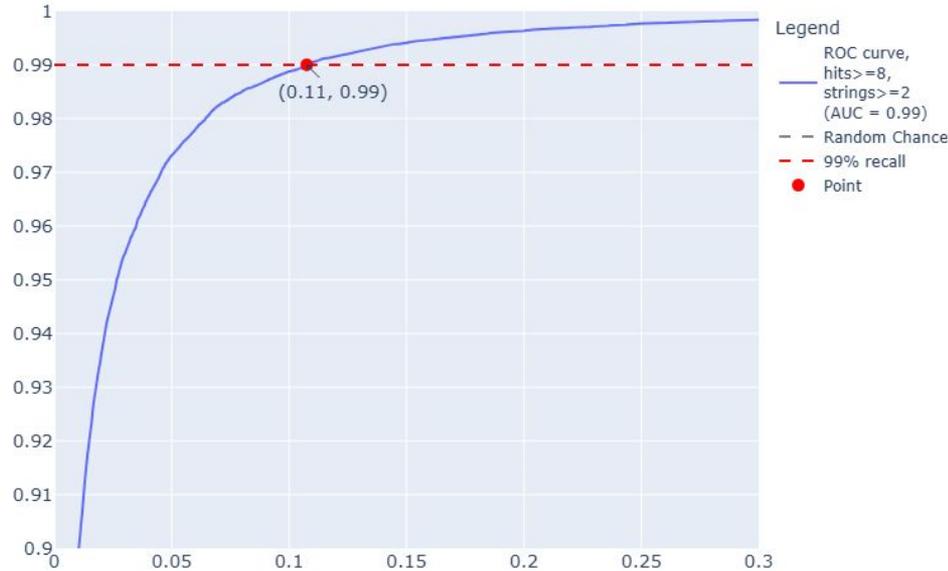
Применение к реальным данным

- Обрезание сигнала на 100 ф.э. (учет насыщения детекторов)
- Добавление 1,5х ожидаемых ошибок показаний ОМ (учет систематики, в том числе МК)
- Флуктуации интегрального заряда, 10% от зарегистрированного



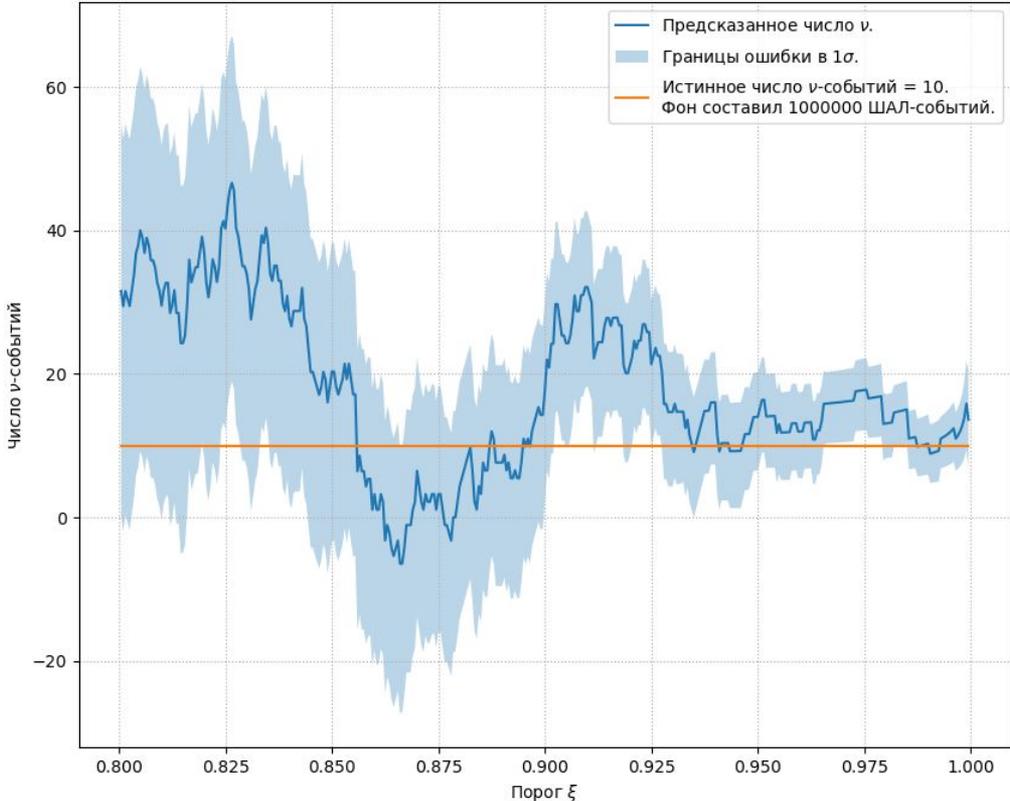
Подавление фона ШАЛ

ROC AUC Curve



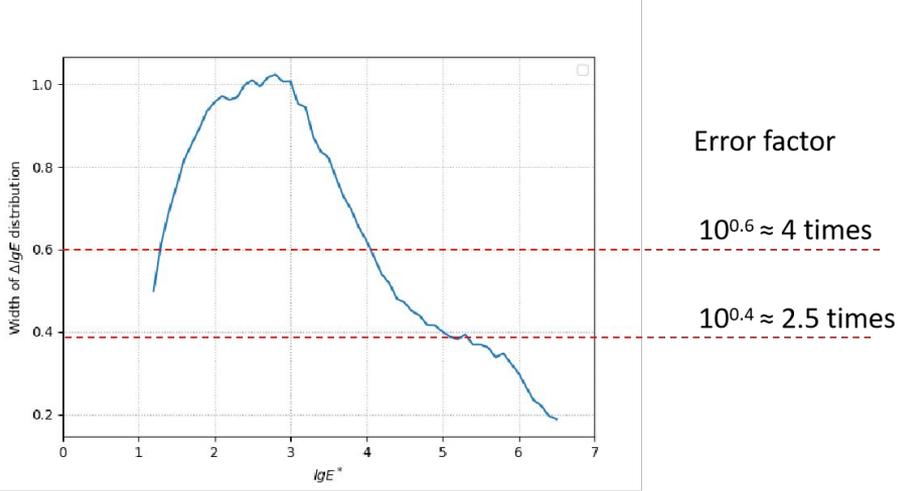
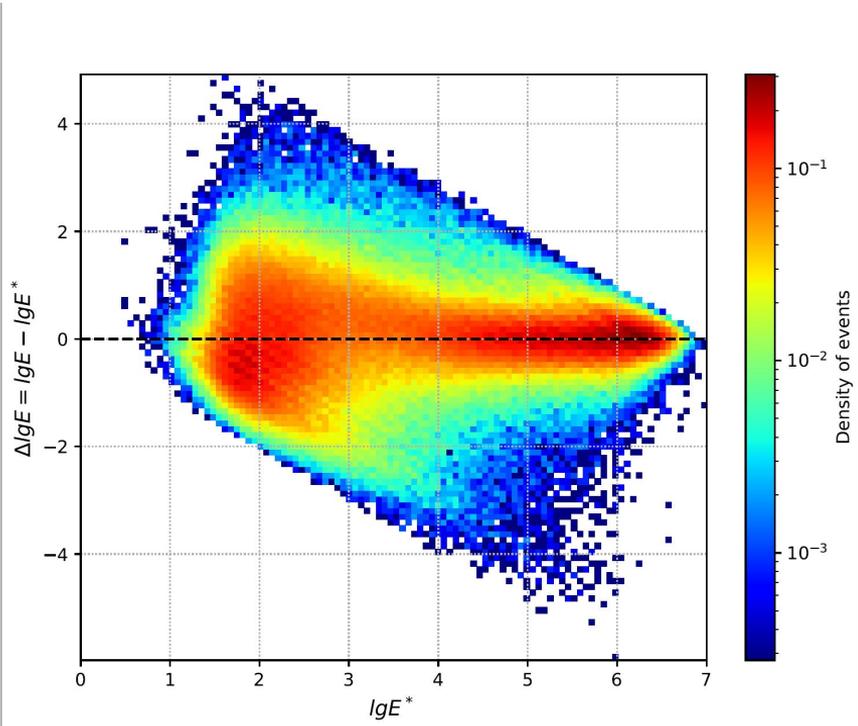
- Скорость обработки данных быстрее до 10^3+ раз.
- 90% ШАЛ сразу отклоняется, без необходимости делать реконструкцию.

Оценка потока нейтрино



Можно обнаружить 10 событий с ошибкой ± 3 на фоне 1 миллиона шумовых событий.

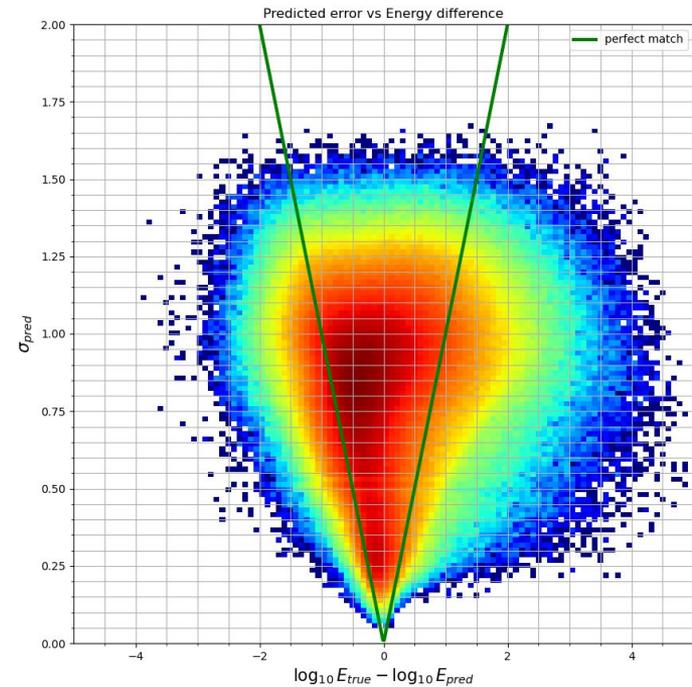
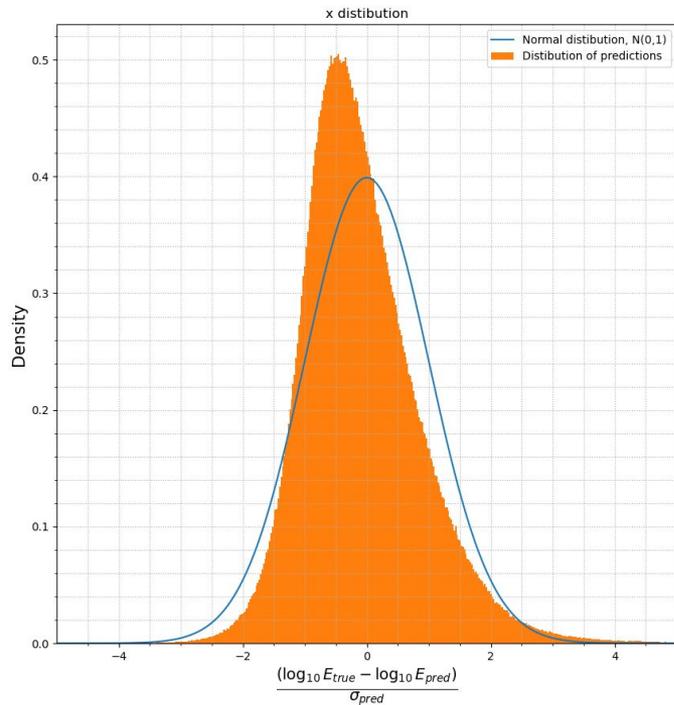
Реконструкция энергии



(Энергия в ГэВ)

Оценка неопределенности энергии

$$\text{Loss} = L_{\text{reco}} + \ln \sigma^2 + L_{\text{reco}} / \sigma^2$$



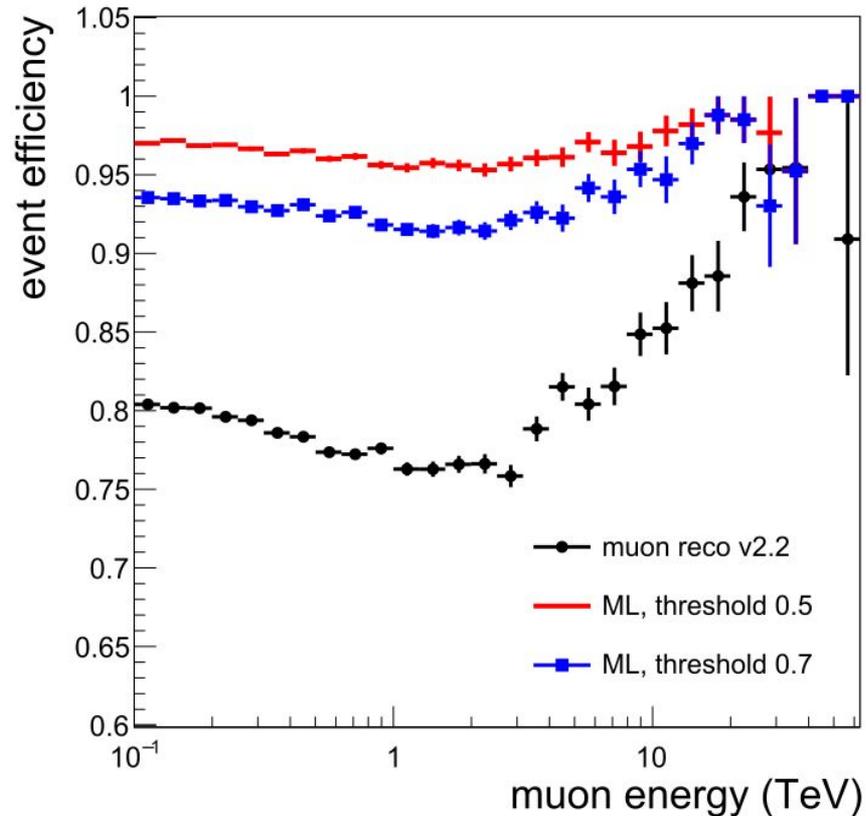
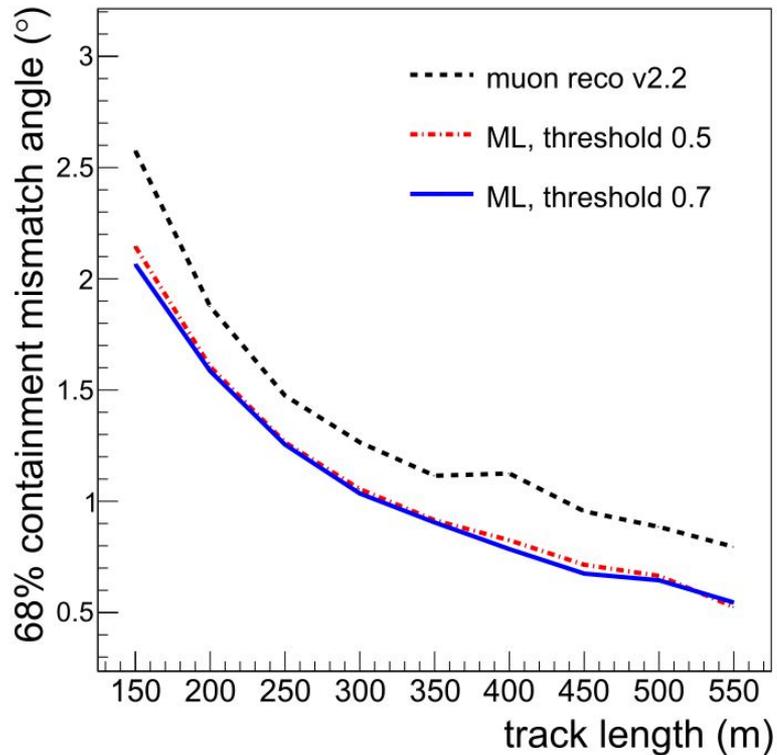
Заключение и перспективы

- Внедрение нейронных сетей позволит существенно увеличить скорость обработки данных
- Метрики нейронных сетей сравнимы или лучше своих алгоритмических аналогов
- Можно разработать цепочку анализа, полностью основанную на машинном обучении

- Основная проблема - показать возможность применения нейронных сетей к реальным данным

Appendix

Влияние на качество реконструкции



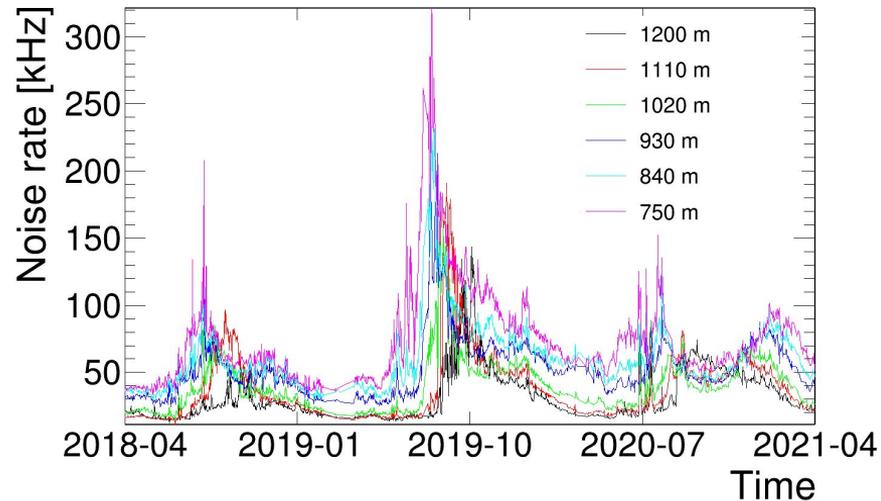
Пример: подавление шумовых срабатываний ОМ

Мотивация:

Шумовые срабатывания составляют **~85%** от собранных данных, уровень сигнала **~1-3 ф.э.**

Кат на уровень сигнала был бы слишком сильным: уберет сигнальные хиты, важные для реконструкции.

Нужно убирать как можно больше шума с наименьшей возможной ошибкой.



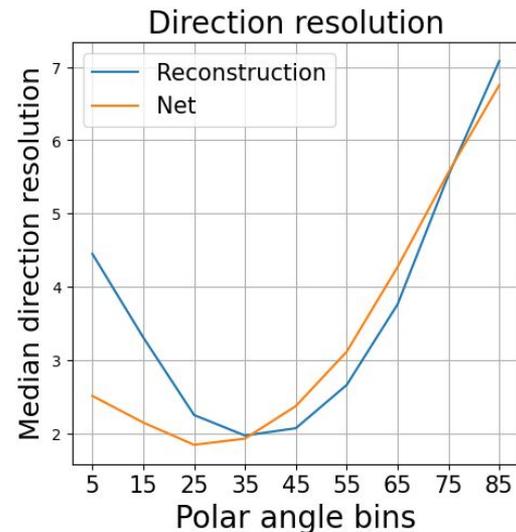
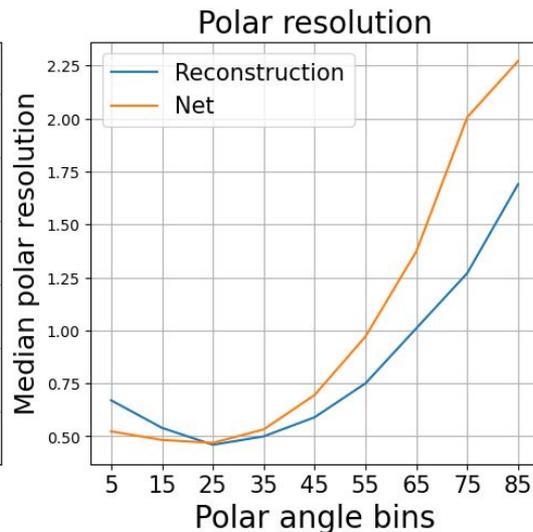
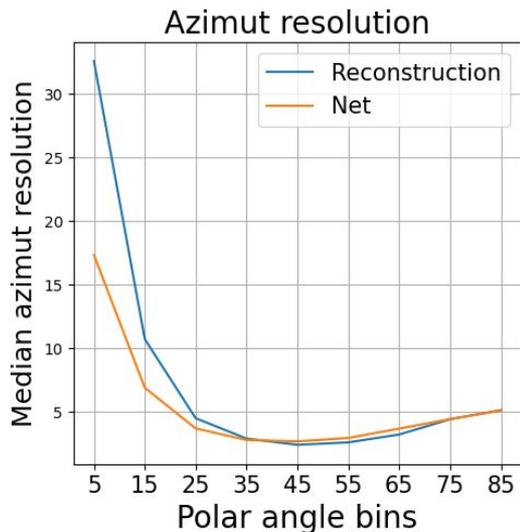
Итоговые метрики

Лучшие метрики показала EdgeGNN (1801.07829)
(введение ребер с информацией помогает нейросети)

Standart Reconstruction			
Metrics	Azimuth Angle	Polar Angle	Direction
50% Resolution	5.42	0.53	2.62
68% Resolution	13.2	0.95	5.62

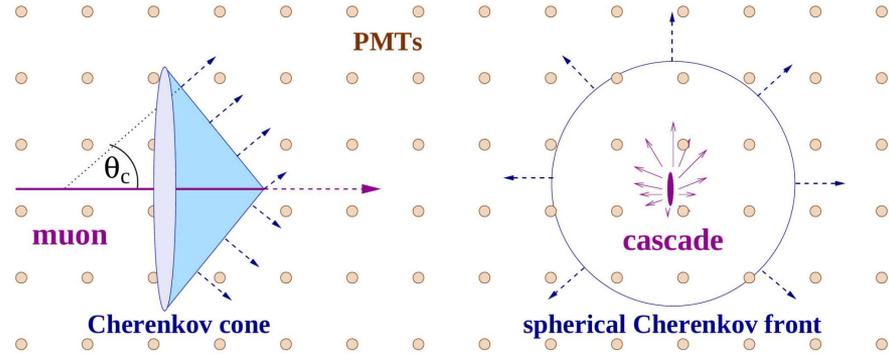
GCN Net			
Metrics	Azimuth Angle	Polar Angle	Direction
50% Resolution	4.16	0.53	2.10
68% Resolution	7.13	0.83	3.18

CNN Net			
Metrics	Azimuth Angle	Polar Angle	Direction
50% Resolution	4.77	0.63	2.57
68% Resolution	7.90	0.93	3.62



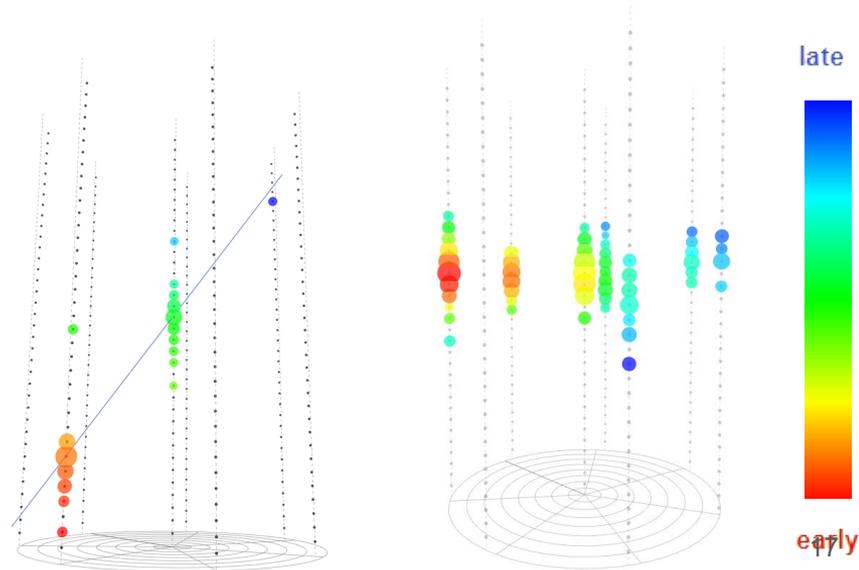
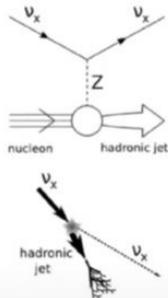
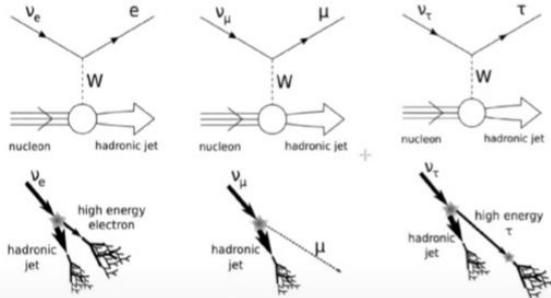
Байкал-GVD

Трэковые и каскадные события:



“Заряженные токи”: взаимодействие через W^\pm бозоны

“Нейтральный ток”:
Взаимодействие через Z^0 -бозон



Основы нейронных сетей



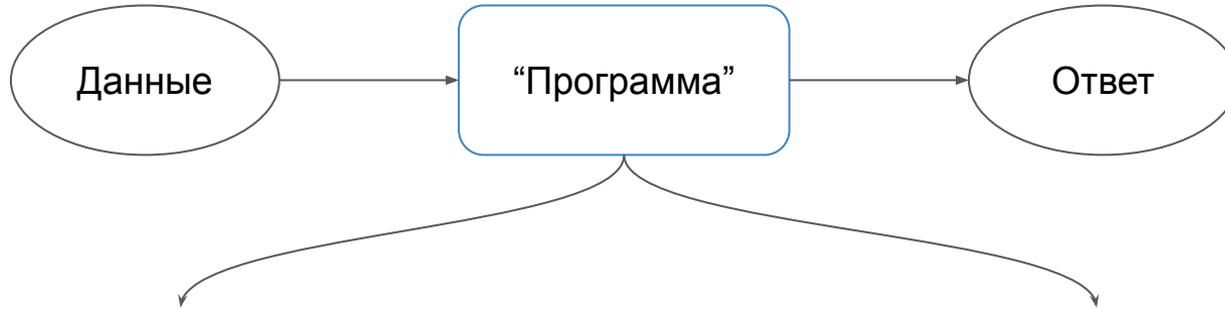
Стандартные алгоритмы:

“Программа” - **фиксированный**,
написанный человеком **алгоритм**
решения задачи.

Как выделить признаки?



Основы нейронных сетей



Стандартные алгоритмы:

“Программа” - **фиксированный**, написанный человеком **алгоритм** решения задачи.

Как выделить признаки?

Машинное обучение:

“Программа” - **обучающийся на примерах** алгоритм выделения оптимальных признаков.

Программы, создающие оптимальные алгоритмы