

Применение методов машинного обучения в эксперименте Байкал-GVD



План доклада

1. Введение:

- a. Основы нейронных сетей
- b. Эксперимент Байкал-GVD

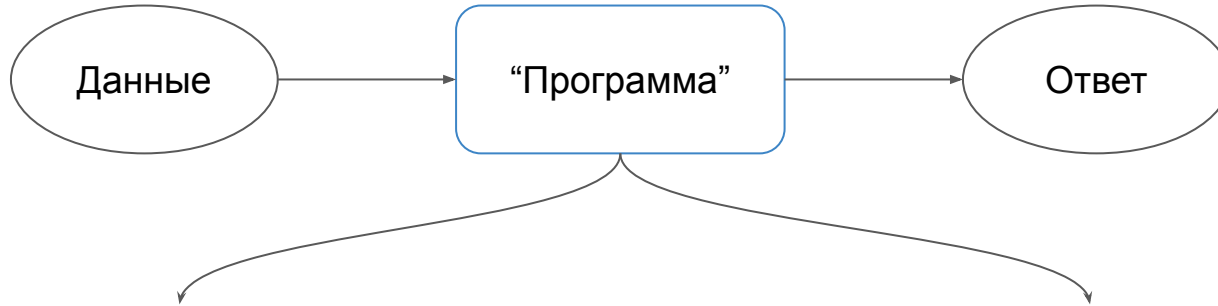
2. Нейронные сети для:

- a. Подавления шумовых срабатываний детекторов
- b. Разделения нейтринных и мюонных событий
- c. Реконструкции угла прилета нейтрино



I. Основы нейронных сетей

Основы нейронных сетей



Стандартные алгоритмы:

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

Как выделить признаки?

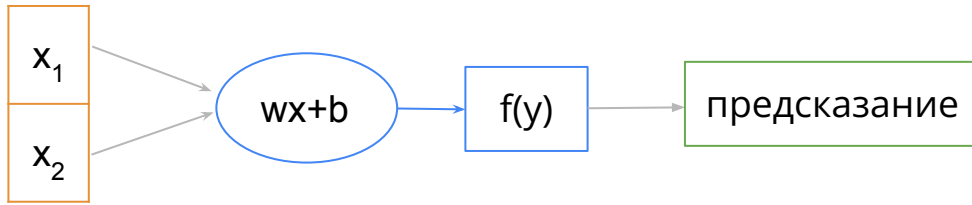
Машинное обучение:

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

Программы, создающие оптимальные алгоритмы

Основы нейронных сетей

Пример: по координате точки на плоскости предсказать, лежит ли она внутри окружности единичного радиуса.



Предсказание:

$[0;1]$, 0 - внутри, 1 - снаружи

Мера ошибки:

штраф = $|\text{предсказание} - \text{правда}| \geq 0$

Инициализация:

Выбираем случайно w и b

Оптимизация:

Пока возможно улучшение:

1. Для заданной точки, считаем предсказание
2. Считаем величину функции штрафа
3. Методом градиентного спуска изменяем w и b , чтобы минимизировать штраф.
4. Берем следующий “обучающий пример”

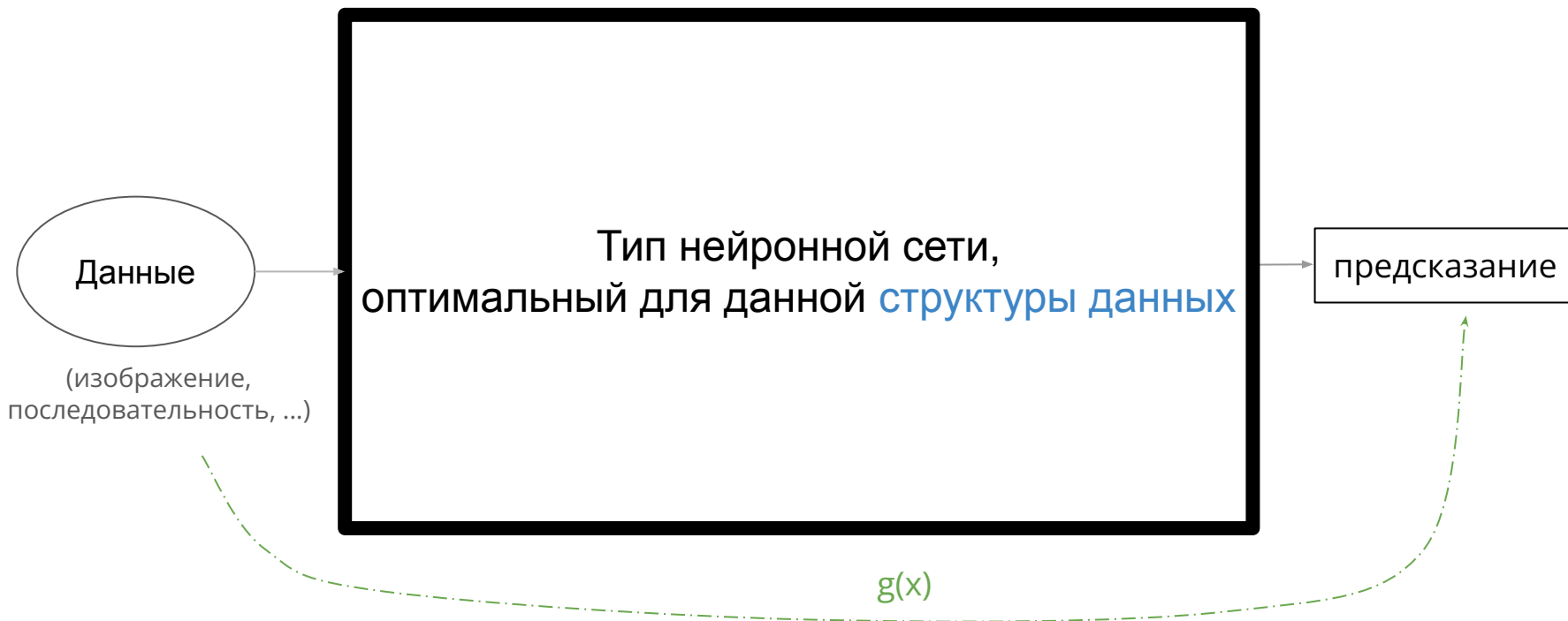
Итог:

Оптимальные значения w и b .

Найденные оптимальные параметры и есть итоговая “программа”.

Сила машинного обучения

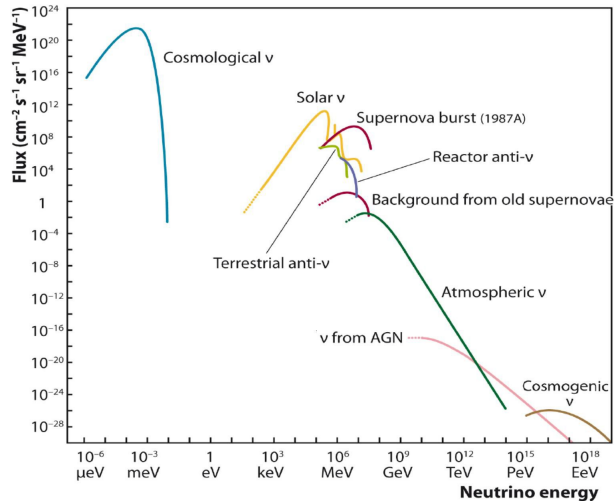
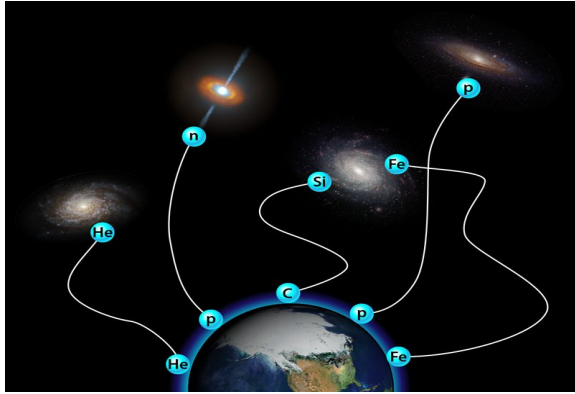
Нейросеть способна
аппроксимировать
любую функцию



II. Baikal-GVD

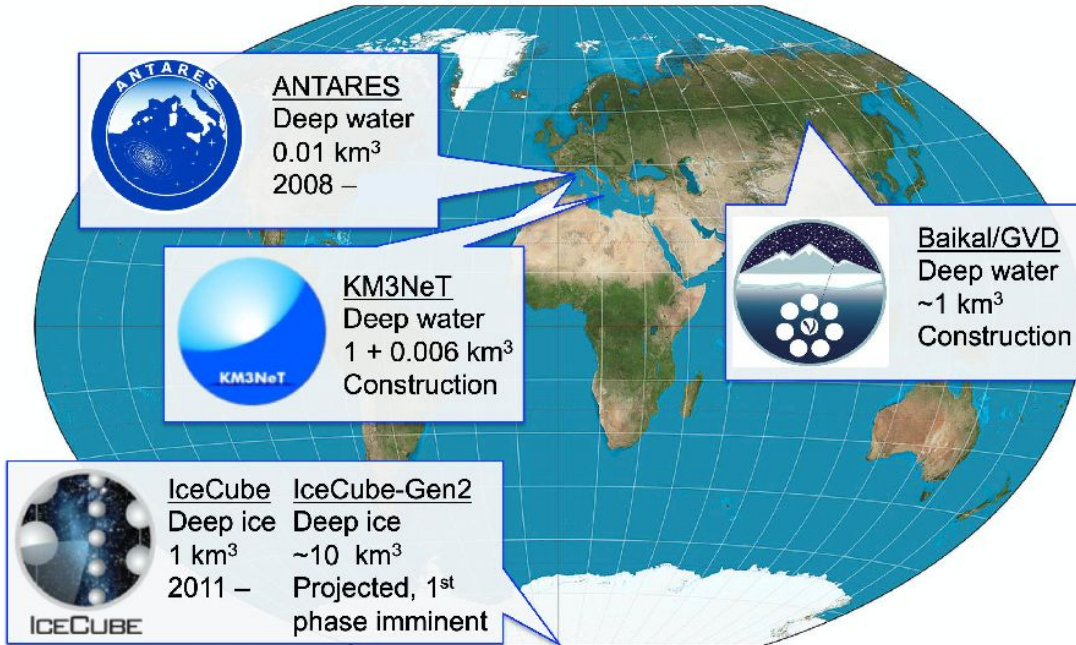


Байкал-GVD: интересность нейтрино



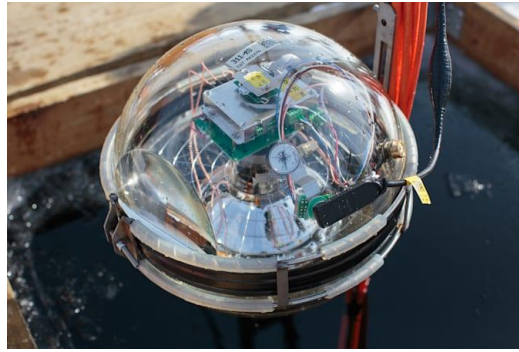
- Напрямую указывают на источник
⇒ проверка моделей
- Вселенная почти прозрачна для ν
- Многоканальная астрономия
(пример: [не]регистрация нейтрино
вместо с фотонами - [лептонные]
адронные процессы)

Байкал-GVD

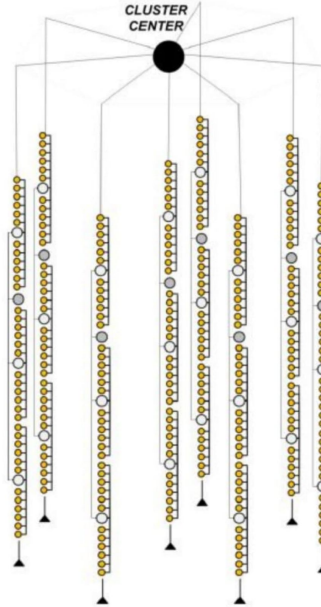
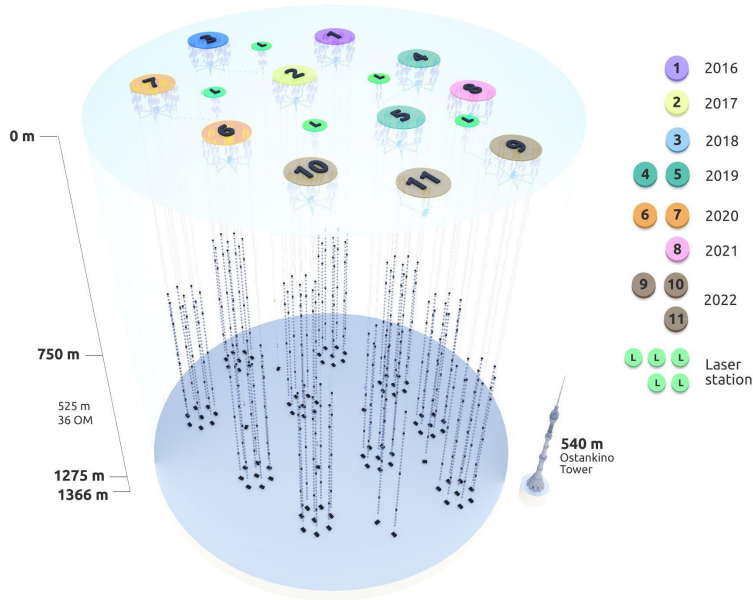


- + Один из 3х больших нейтринных телескопов (по сост. на 23 год, ~0.5 км³)
- + Дополняет IceCube: вместе покрывают оба полушария
- + Эффективная эксплуатация:
 - : вблизи населенных пунктов
 - : оборудования можно ставить пока озеро подо льдом
- + Возможно сотрудничество с Китаем для расширения эксперимента
- Не чувствителен к нейтринным осцилляциям (ORCA@KM3NeT)

Байкал-GVD



Детектор \Rightarrow 11 кластров \Rightarrow
 \Rightarrow 8 струн \Rightarrow 36 оптических
модулей (ОМ)



Живая калибровка положений
ОМ (точность ~ 20 см)

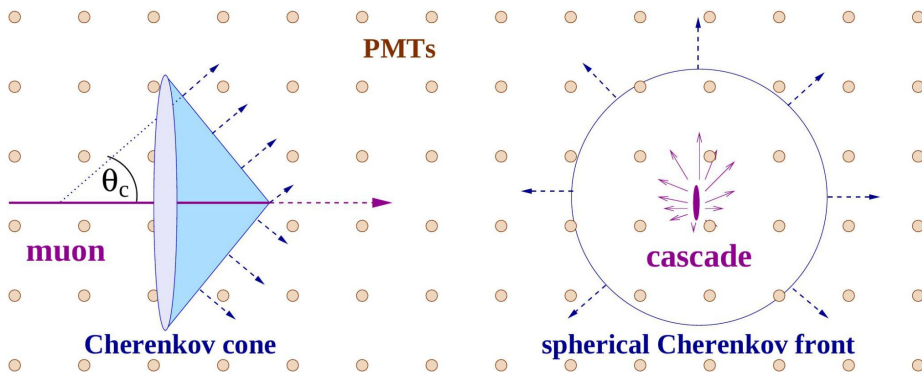
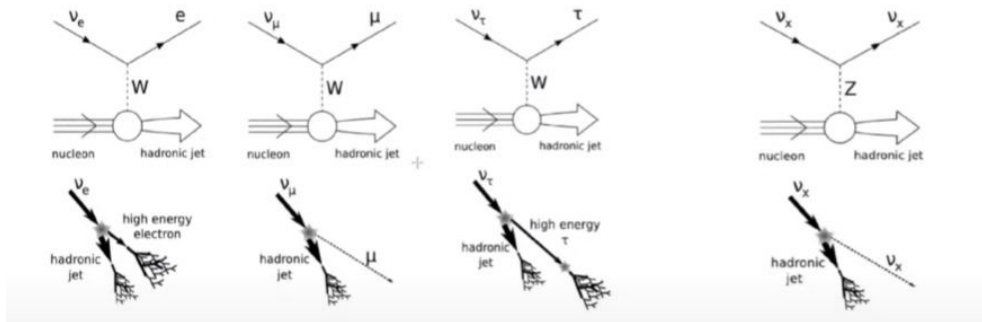
Точность определения
времени срабатываний ~ 2 нс.

Байкал-GVD

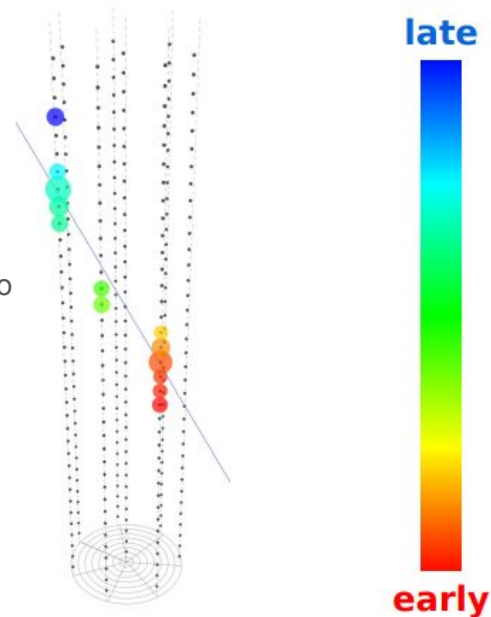
ОМ регистрируют фотоны от череновского излучения (+ шум)

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

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



Пример трэкового события





III. Нейронные сети в Байкал-GVD

Общий план применения нейронных сетей

Считываемые данные



Подавление шумов

→ стандартная
реконструкция



Выделение нейтринных событий

→ оценка потока нейтрино



Предварительная \oplus
реконструкция →

Реконструкция параметров событий
(энергия, направление)

Используемое Монте-Карло

Два типа событий:

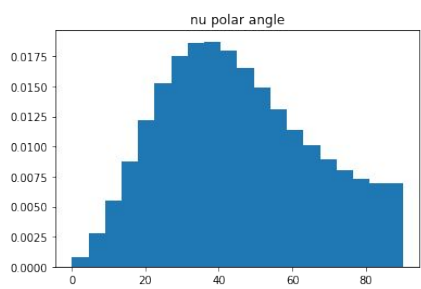
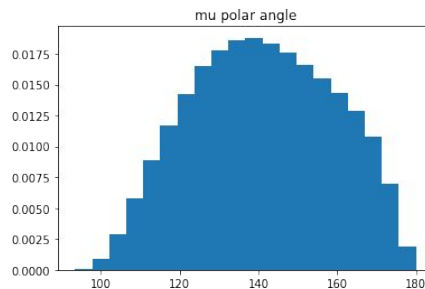
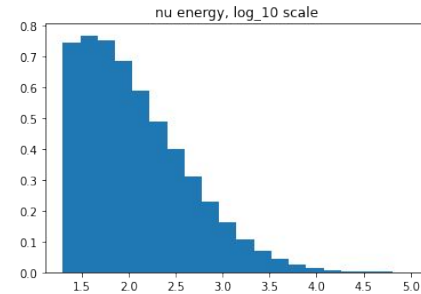
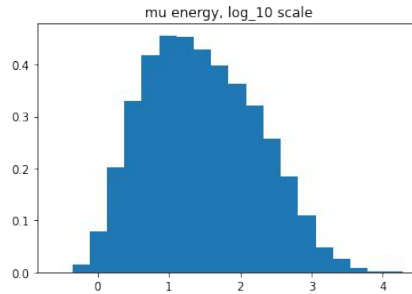
- 1) Мюоны от космических лучей (μ)
- 2) ν_{μ} -индуцированные события

Полная симуляция развития ливней и распространение частиц в воде.

Ограничение:

Один кластер, фиксированная геометрия

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; источник сигнала (какой μ) и его тип (трэк/каскад)

1. Подавление шумовых срабатываний ОМ

Мотивация:

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

Кат на уровень сигнала слишком сильный: уберет сигнальные хиты, важные для реконструкции.

Нужно убирать как можно больше шума с наименьшей возможной ошибкой.

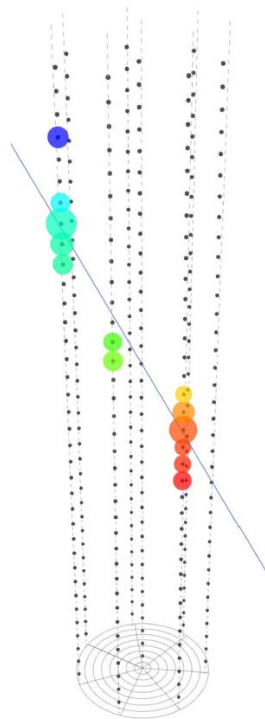
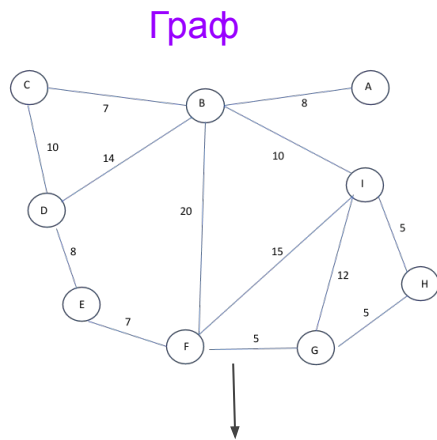
Два режима работы нейросетей:

А) “Полный”: пытаемся идентифицировать все сигнальные хиты

Б) “Трэковый”: ищем сигнальными хиты только с $t_{res} < 20ns$

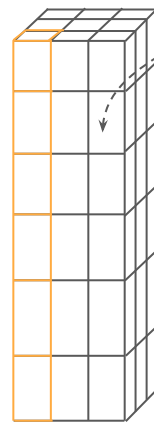
Представление данных

Несколько вариантов представления данных.



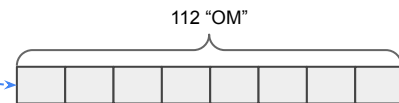
Геометрическое

Показания детекторов



3D свертки

Причинное (временное)



Упорядочивание по времени

1D свертки и рекуррентные сети

Используемая нейросеть

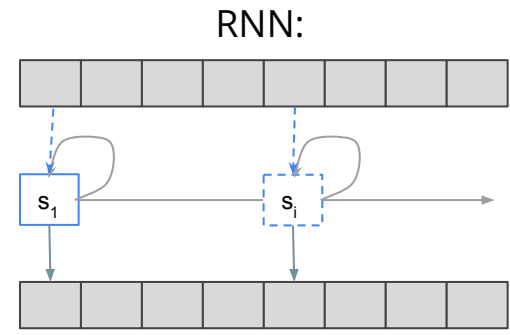
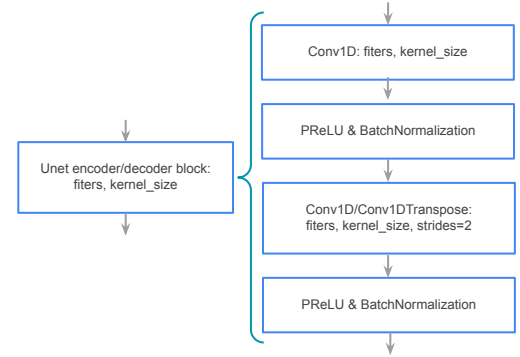
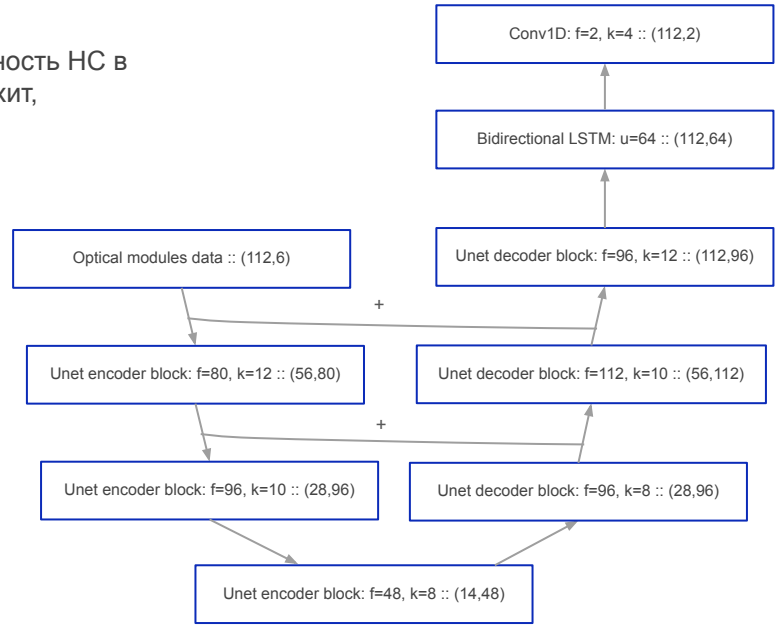
Входные параметры:

- 1-3) координаты x, y, z
- 4) интегральный сигнал
- 5) время активации
- 6) маска
(вспомогательный/настоящий хит)

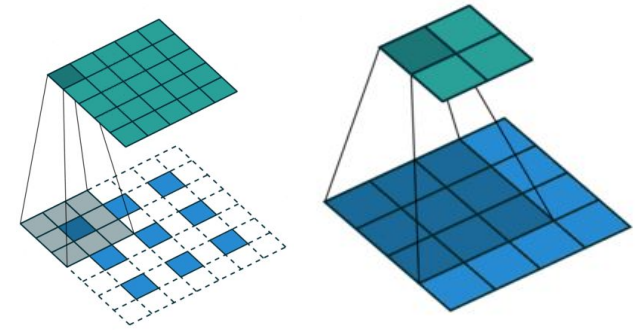
Выход:

Для каждого ОМ, уверенность НС в том, что это сигнальный хит, $p \in [0;1]$

Unet: изучение локальных и глобальных свойств (горизонтальные связи и вертикальные связи)



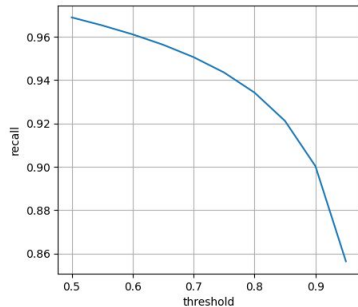
CNN:



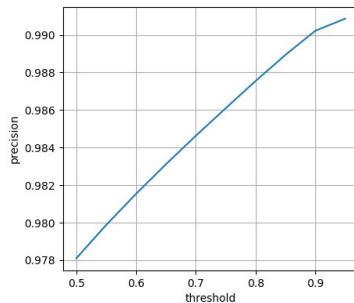
Vincent Dumoulin, Francesco Visin - A guide to convolution arithmetic for deep learning

1. Метрики

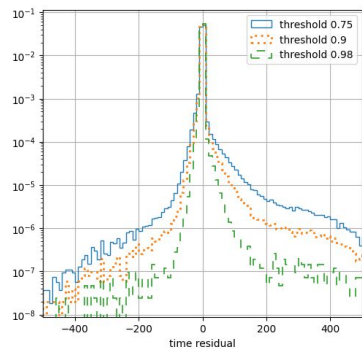
Точность



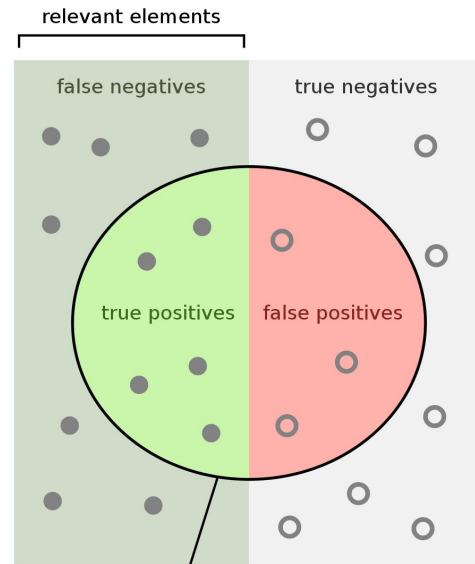
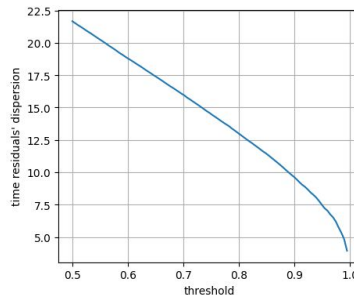
Полнота



Распределение t_{res}



t_{res}



retrieved elements

How many retrieved items are relevant?

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

How many relevant items are retrieved?

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

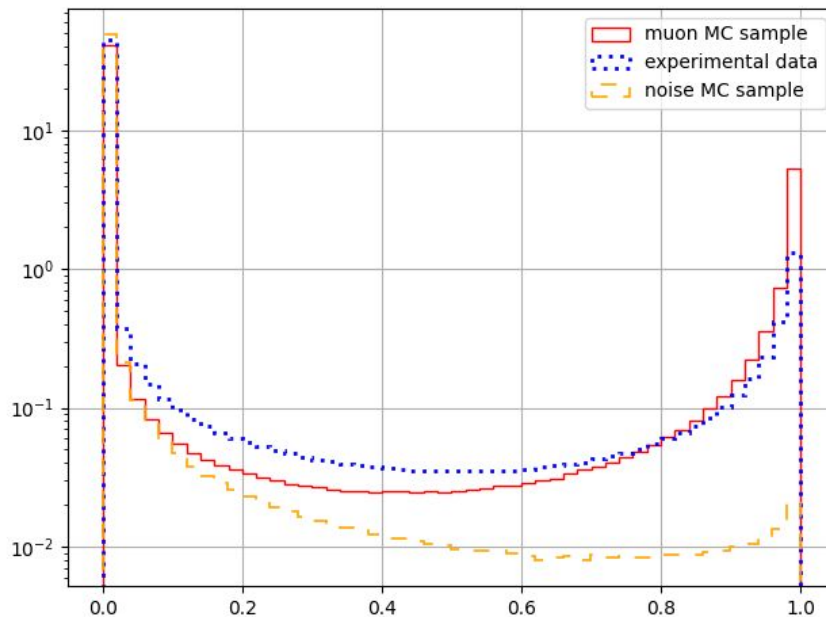
Точность (precision) и полнота (recall) лучше на 4% и 1% соответственно!

Скорость обработки быстрее до 10^3+ раз.

1. Применение к реальным данным

Модулирование данных:

- 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



Плохое согласие между МК и реальными данными: недооценка шума и сигнальных мюонов.

Ожидается заметно лучше согласие при переходе к новому МК (в процессе).

2. Разделение μ - и ν_μ -индуцированных событий

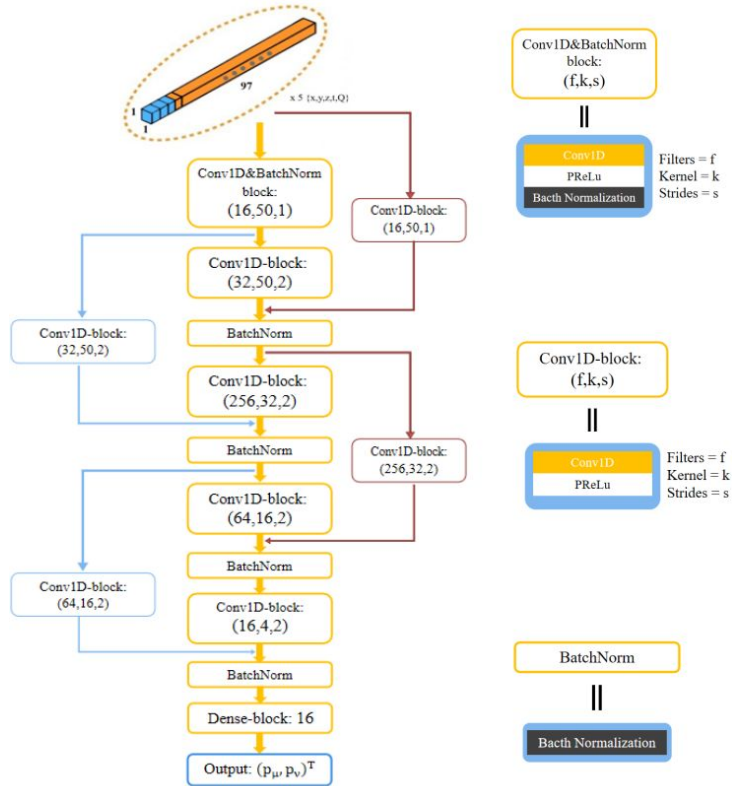
Мотивация:

Отношение μ к ν_μ событиям порядка 10^6 - 10^7 .

Стандартные каты достаточно жесткие.

Можно добиться лучшего разделения с более слабыми катами.

2. Разделение μ - и ν_{μ} -индуцированных событий



Специальная функция потерь:
focal loss (arXiv:1708.02002)

$$\text{Loss} = (1 - p_{\text{correct}})^2 * \text{все.}$$

Гарантирует малое количество ложноположительных identifications (ЛПИ).

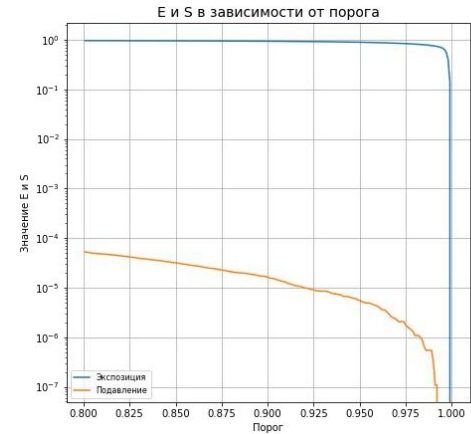
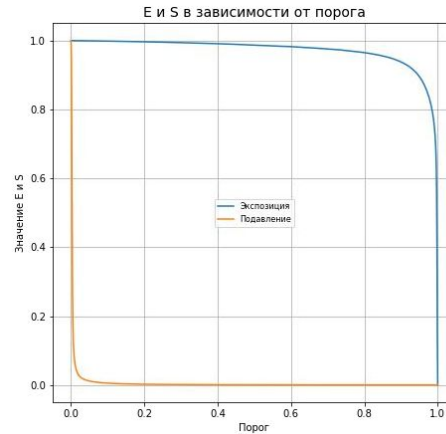
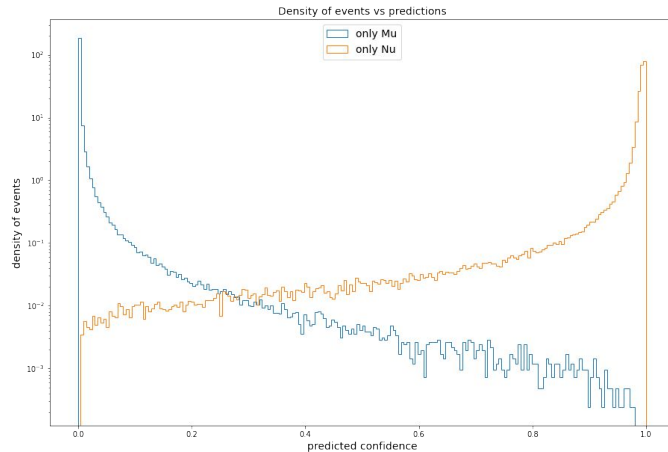
Дополнительно:

Тренировка 3х сетей и усреднение их предсказаний для ещё большего уменьшения ЛПИ.
("коллегия экспертов")

2. Метрики

Экспозиция $E()$ - доля нейтрино справа от порога классификации

Подавление $S()$ - доля мюонов справа от порога классификации



Значение порога классификации определяется целеполаганием:
ЛПИ $< 10^{-n}$, или наименьшая ошибка оценки потока нейтрино

2. Оценка потока нейтрино

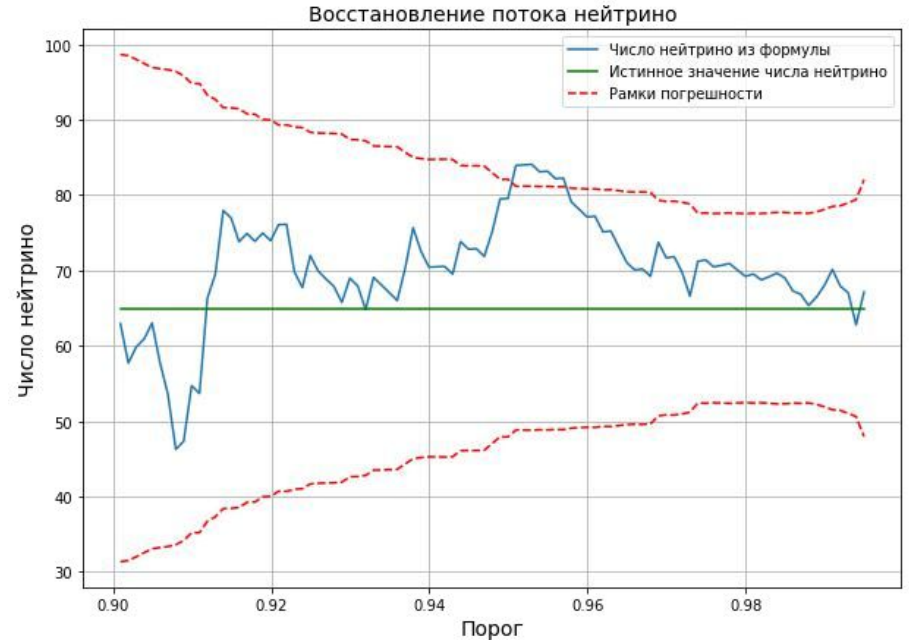
$$n_{\nu}^i \approx \frac{n^i(\xi) - S^0(\xi)n^i}{E^0(\xi) - S^0(\xi)}$$

S^0 , E^0 - подавление и экспозиция, оцененные на тестовом МК наборе данных.

ξ - порог классификации.

$n^i(\xi)$ - количество нейтринных событий, идентифицированных сетью правее порога.

Можно оценить ошибку!

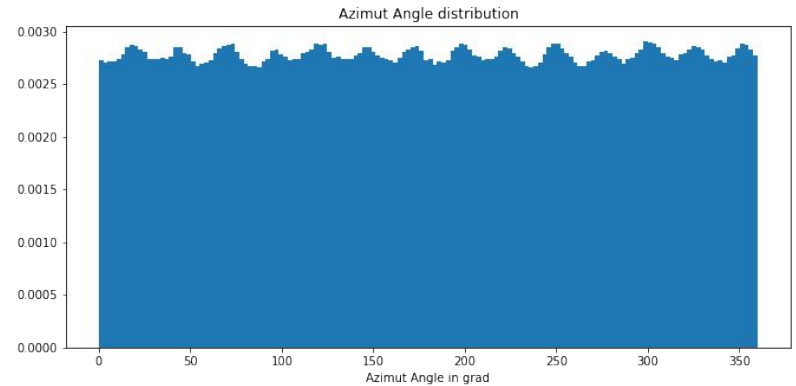
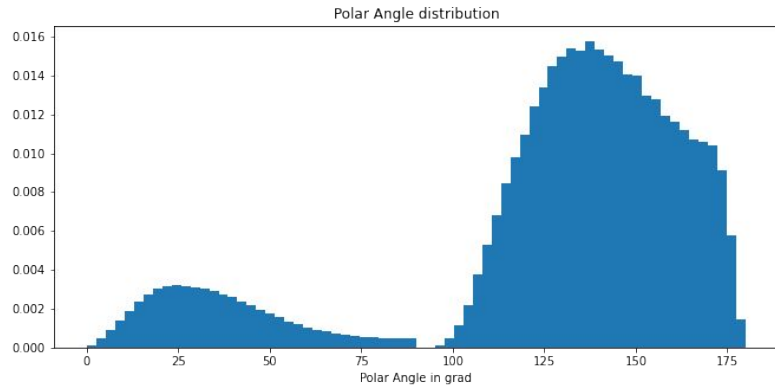


Отношение μ к ν_{μ} : $\sim 500\,000$

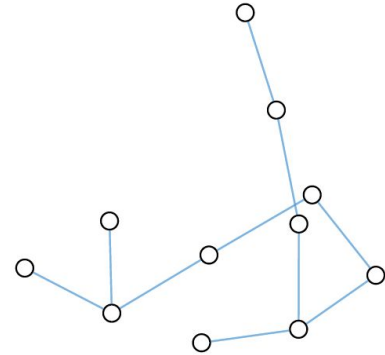
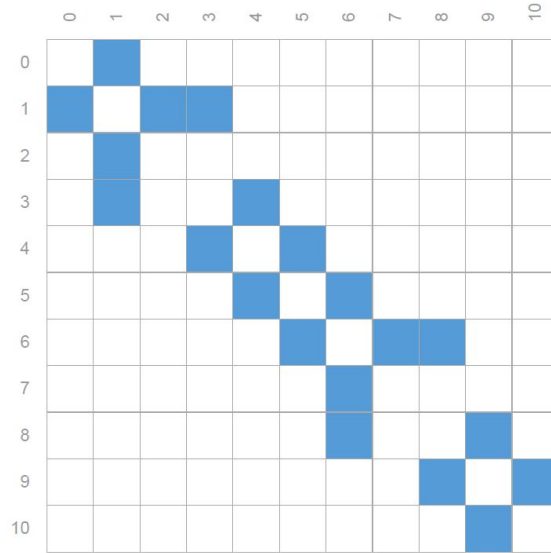
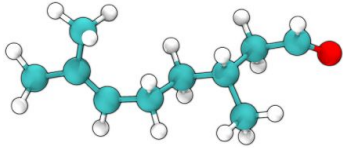
3. Реконструкция угла прилета нейтрино

Мотивация:

Улучшение углового разрешения телескопа

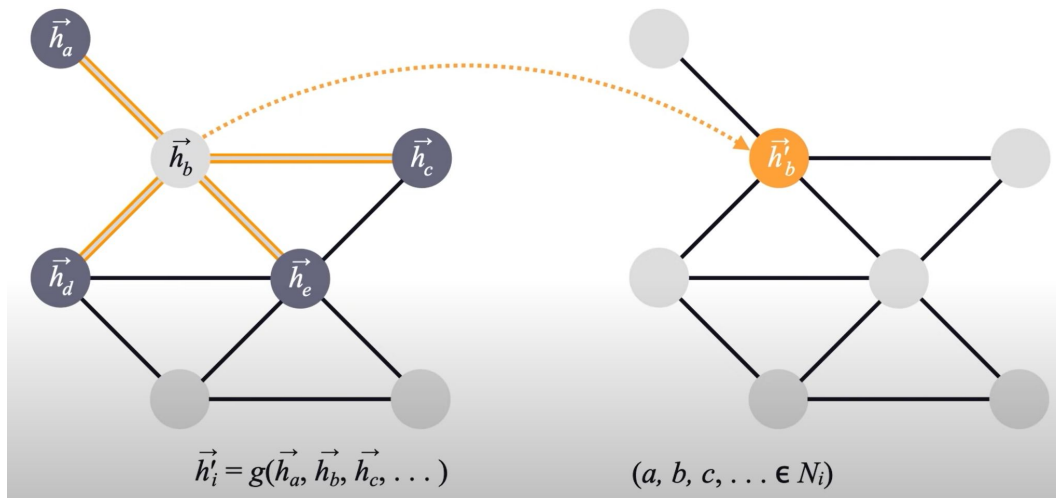


3. Графовые сети



Графы позволяют учитывать более сложные связи в данных

3. Графовые сети



Графовые сети “обновляют” граф: на следующем шаге значение вершины является функцией от а) её соседей, б) связывающих рёбер, в) глобальных агрегированных свойств графа.

Для реконструкции угла прилета, агрегируется информация со всего графа.

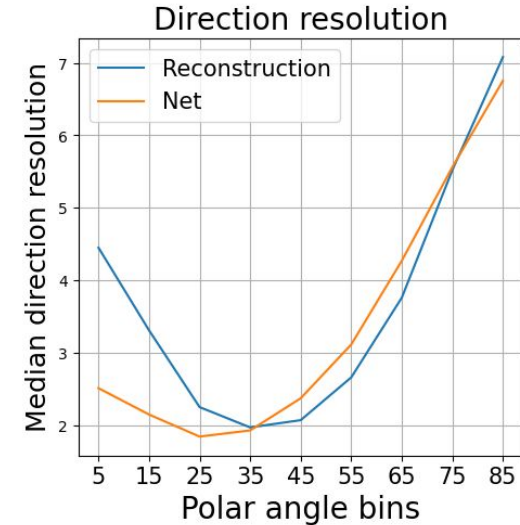
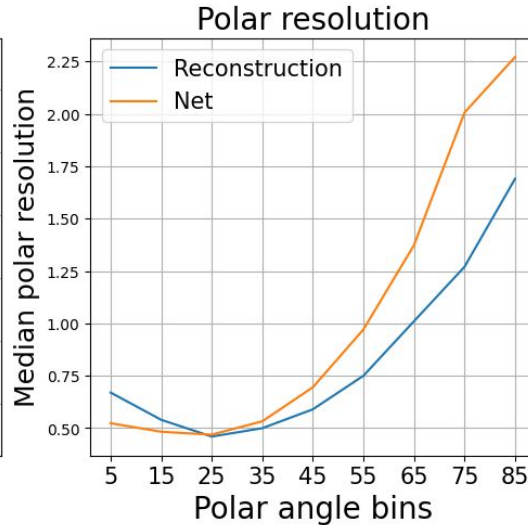
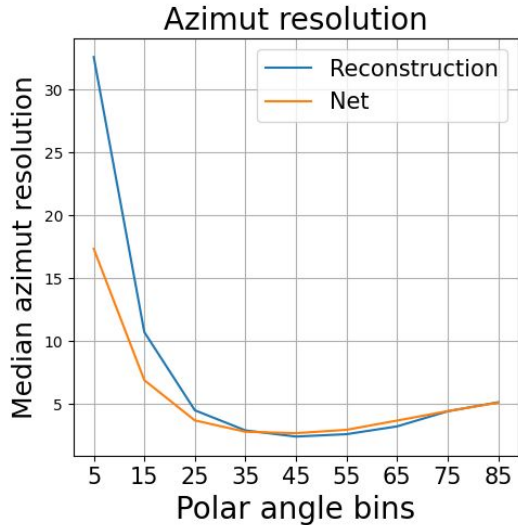
3. Итоговые метрики

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

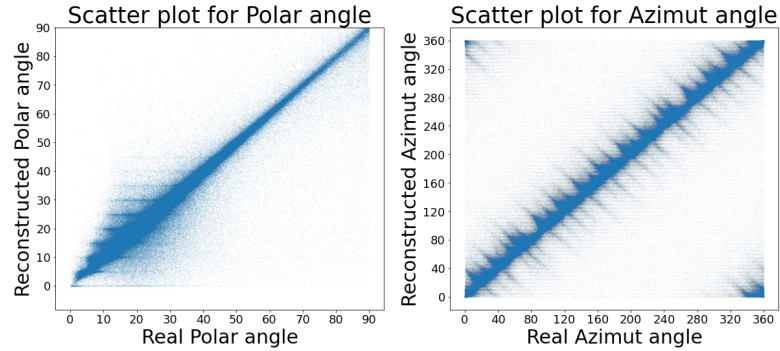


Заключение

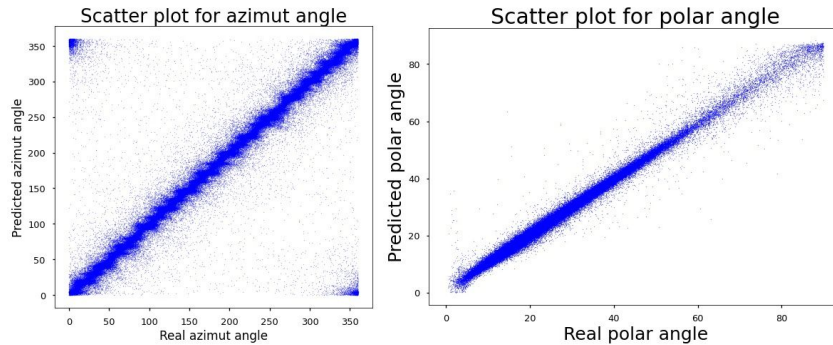
- Нейронные сети позволяют улучшить качество реконструкции событий в эксперименте Байкал-GVD
- Ожидается улучшение согласованности МК и реальных данных при переходе к новой версии МК
- Планируется разработка новых нейронных сетей для реконструкции энергии, выделении трэковых и каскадных хитов, а также оптимизация существующих сетей

Аппендикс

3. Распределения реконструированных углов



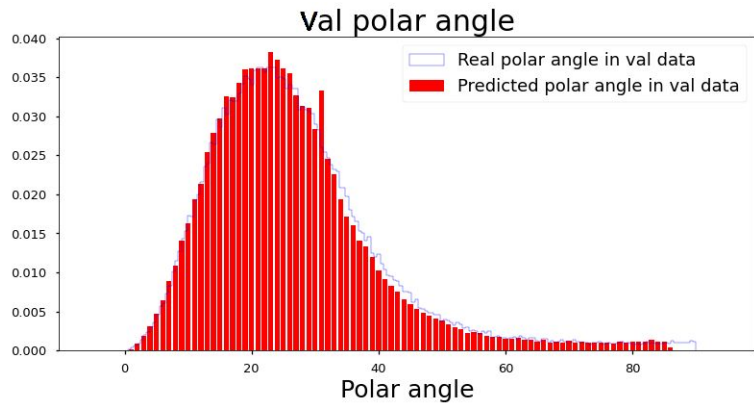
Алгоритмическая реконструкция



Нейросетевая реконструкция

3. Распределения реконструированных углов

Нейросетевая реконструкция



Алгоритмическая реконструкция

