



Федеральное государственное автономное образовательное учреждение высшего образования
«Национальный исследовательский ядерный университет «МИФИ»

РАЗРАБОТКА РЕШЕНИЙ ПО МОДЕЛИРОВАНИЮ И РЕКОНСТРУКЦИИ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ И МАШИННОГО ОБУЧЕНИЯ

Отчет о научно-исследовательской работе

Научный руководитель:

Герценбергер К.В.

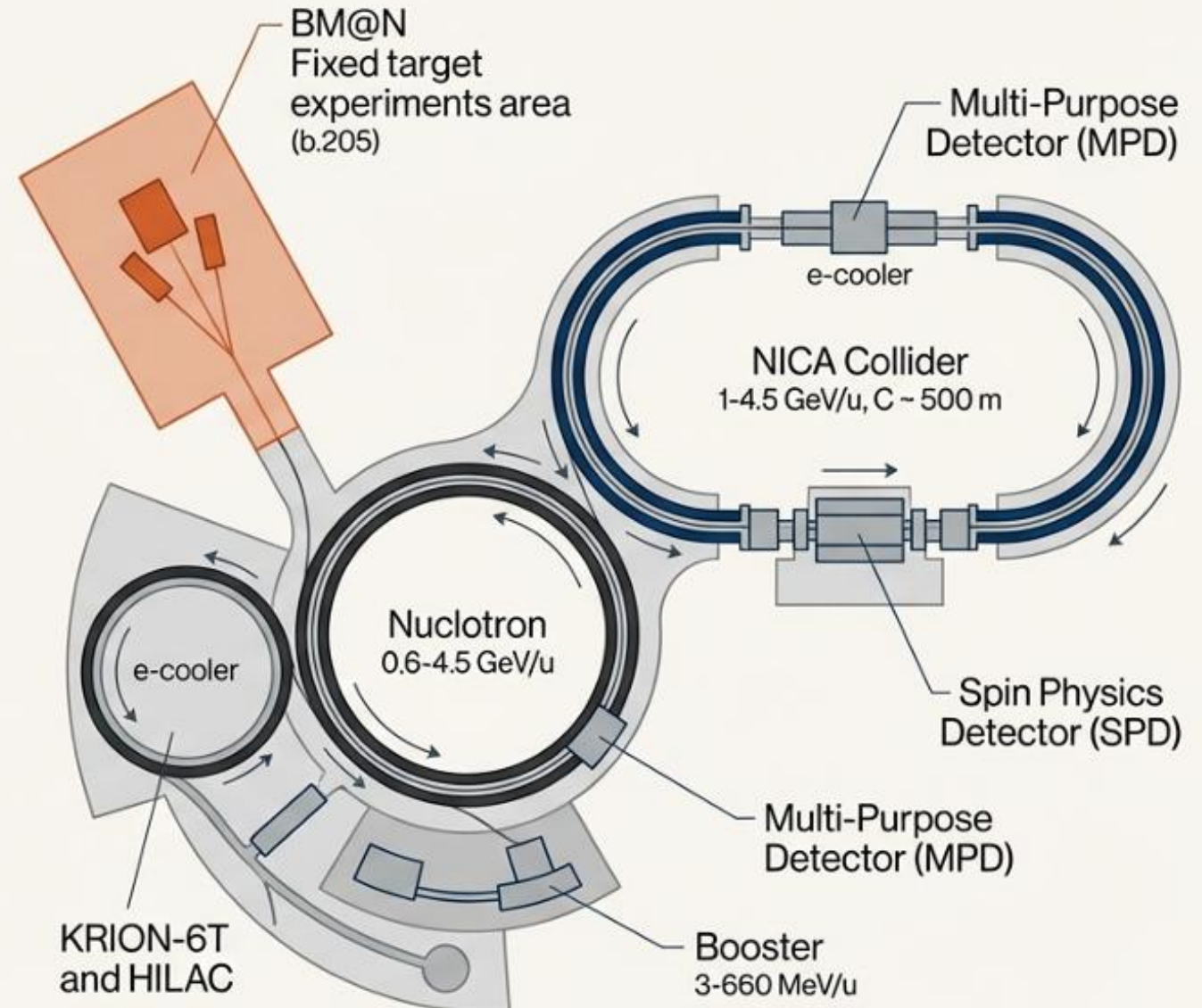
Студент:

Джавадов Д.Н.

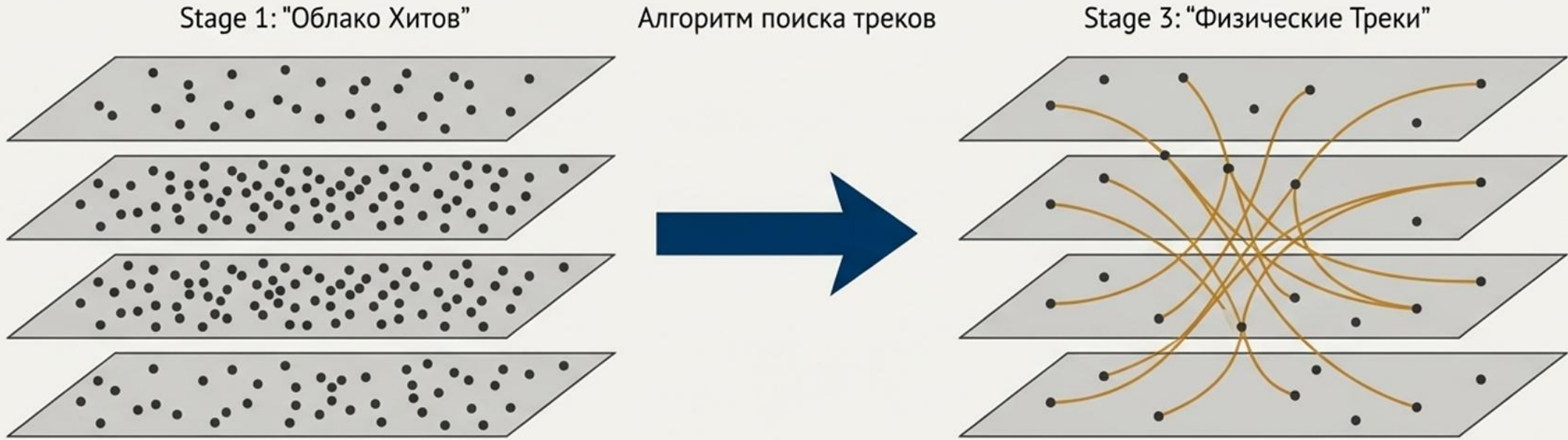
Москва 2025

Комплекс NICA

- Эксперимент BM@N (Baryonic Matter at Nuclotron) ключевой элемент первого этапа проекта NICA (Nuclotron-based Ion Collider fAcility).
- Эксперимент с фиксированной мишенью, изучающий столкновения тяжелых ионов.
- BM@N использует пучки, выведенные из ускорителя, для создания условий экстремальной барионной плотности.



Задача Реконструкции: От Облака Хитов к Физическим Трекам



1. Поиск трек-кандидата:

Самая сложная комбинаторная часть задачи. Необходимо сгруппировать отдельные сигналы (хиты), оставленные частицами на разных слоях детектора, в осмысленные последовательности. На этом этапе происходит основная борьба с комбинаторным взрывом.

2. Фитирование треков

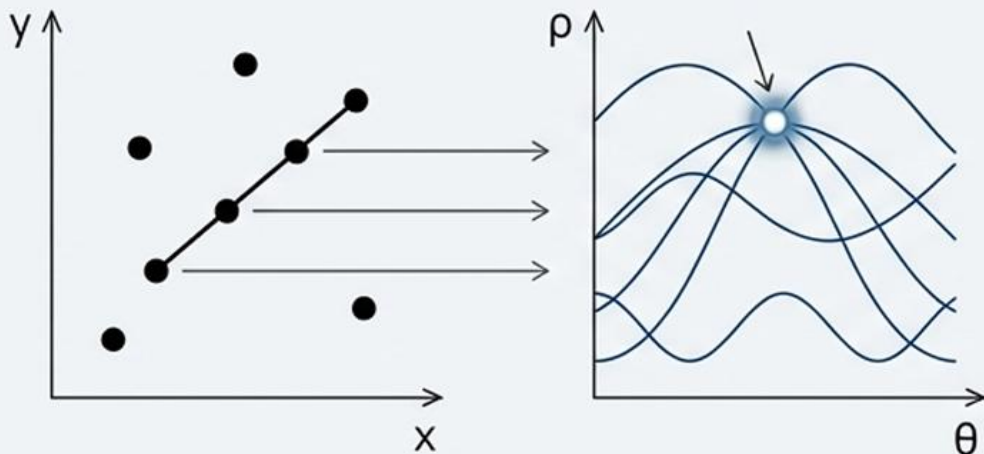
После того как трек-кандидат найден, определяются параметры треков частиц.

Классические подходы к поиску трек-кандидата

В современных экспериментах, включая $BM@N$, для поиска треков применяются алгоритмы, основанные на геометрических и комбинаторных принципах.

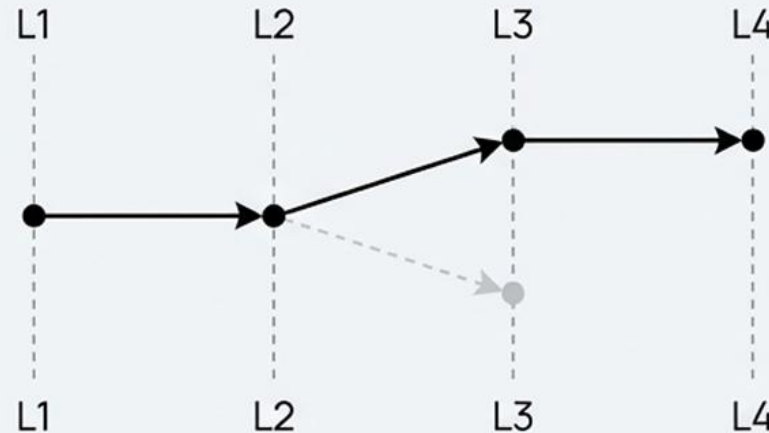
Преобразования Хафа (Hough Transform)

- **Принцип:** Преобразует задачу поиска траекторий в задачу поиска максимумов в пространстве параметров линий
- **Краткое описание:** Каждая точка (хит) «голосует» за все возможные прямые, проходящие через неё, в специальном массиве-аккумуляторе



Клеточные Автоматы (Cellular Automata)

- **Принцип:** Имитирует процесс «выращивания» трека из начальных точек шаг за шагом.
- **Краткое описание:** На каждом шаге алгоритм пытается продлить растущий трек, добавляя подходящие хиты с соседней детекторной плоскости по локальным правилам.



Общий барьер: комбинаторная природа и ее последствия

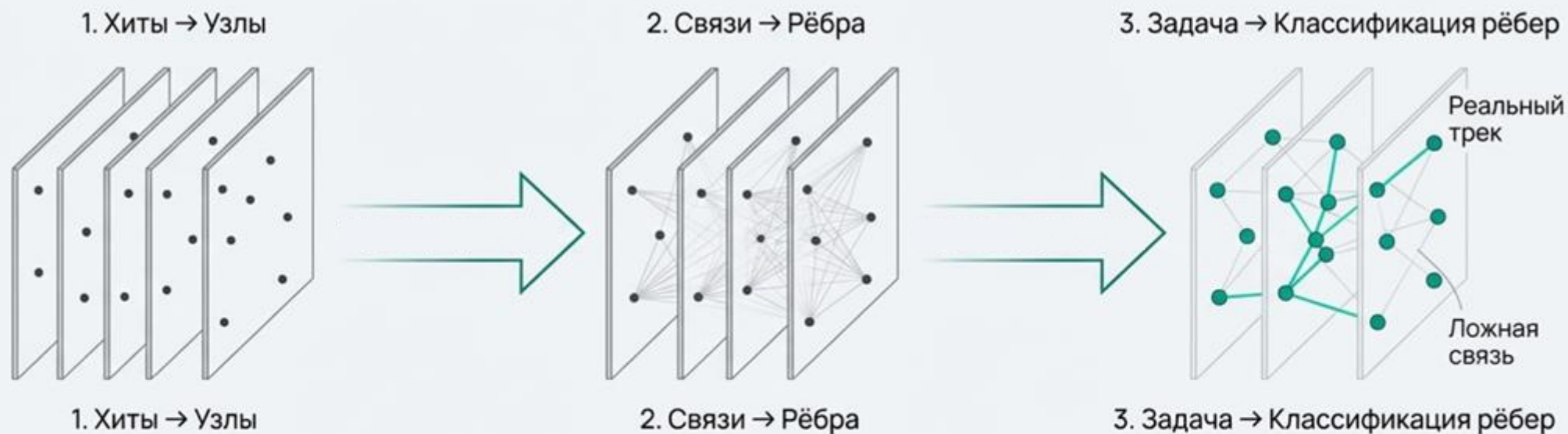
Несмотря на разные подходы, классические методы имеют общее фундаментальное ограничение.

Основной тезис: Высокая плотность фоновых хитов в детекторах $BM@N$ делает перебор возможных комбинаций хитов чрезмерно большим



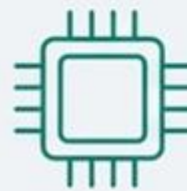
Новая парадигма: представление данных в виде графа

Графовые нейронные сети (GNN) предлагают принципиально иной подход, основанный на естественном представлении данных трекового детектора.



Глобальный анализ

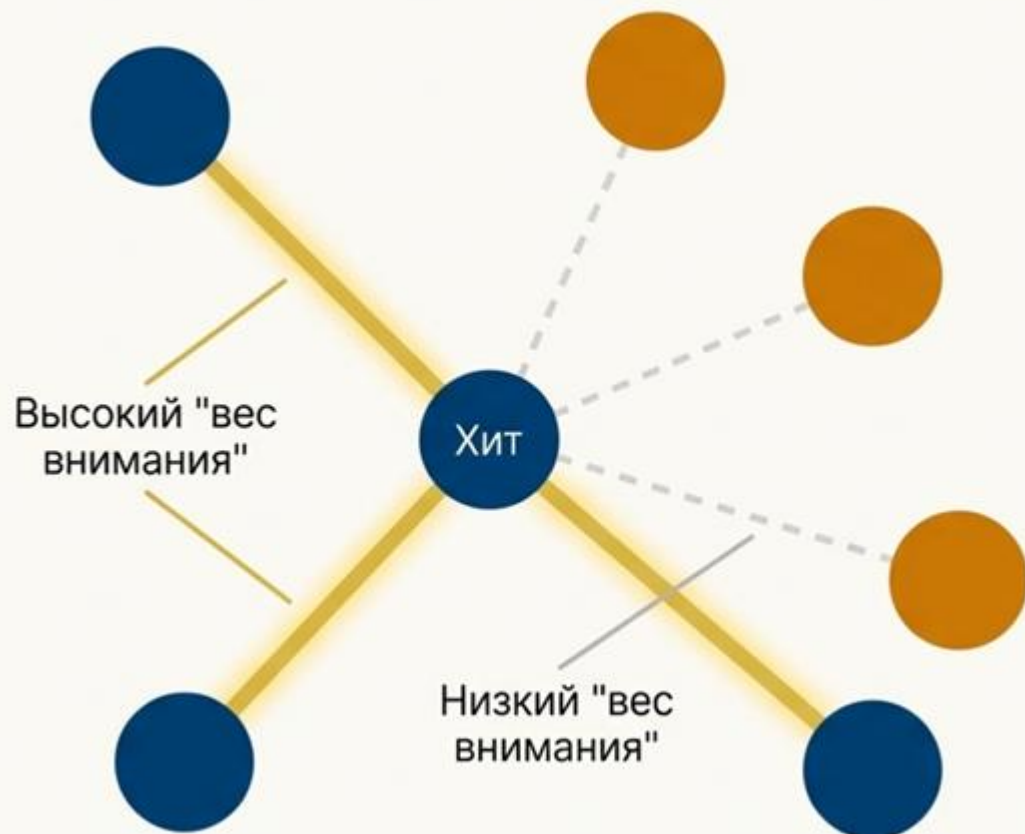
Нейросеть оценивает все взаимосвязи между хитами одновременно в рамках единой структуры.



Масштабируемость

Операции GNN эффективно распараллеливаются, открывая путь к значительному ускорению трекинга.

Инструмент: Graph Attention Network (GAT) – фокус на важном



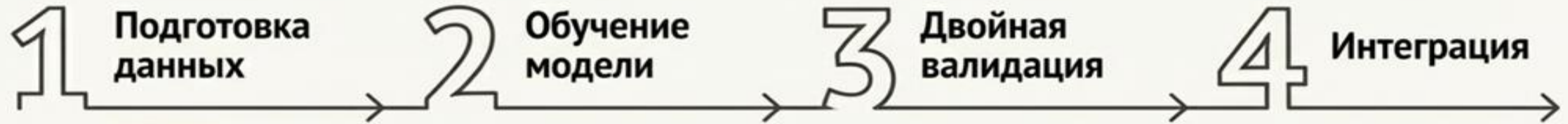
GAT не просто агрегирует информацию, а динамически взвешивает ее.

Механизм внимания (Attention) вычисляет важность каждого ребра (коэффициент внимания)

Модель обучается игнорировать шум и концентрироваться на физических паттернах.

Почему это важно для ВМ@N? В условиях, где ложные хиты превосходят по количеству полезные сигналы, способность GAT к адаптивной фильтрации является решающей для выделения физических траекторий.

Дорожная карта проекта



Подготовка данных

Генерация датасета в BmnRoot, построение графов сначала эвристическим методом k-NN с последующим переходом к физически мотивированному представлению.



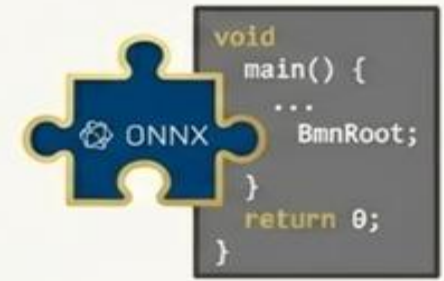
Обучение модели

Реализация и настройка GAT в PyTorch Geometric для бинарной классификации рёбер



Оценка

Оценка по метрикам Машинного Обучения (ML) и Задач Реконструкций



Интеграция

Экспорт обученной модели в формат ONNX для использования в среде C++ (BmnRoot)

Этап 1 – Усложнение графа с учетом физики

Цель: Перейти от абстрактного геометрического графа к физически мотивированному представлению для сокращения ложных гипотез и повышения эффективности обучения.

Планируемые улучшения логики построения графа:

- **Учет магнитного поля:** Введение правил, учитывающих ограничения по кривизне траекторий частицы
- **Физические процессы:** Учет многократного рассеяния частиц
- **Геометрические ограничения:** Последовательность прохождения слоев детектора, минимально допустимые расстояния.
- **Вторичные продукты:** Возможность образования вторичных вершин при распадах частиц.

Ожидаемый результат: Значительно сократить количество ложных рёбер-кандидатов.

До: Геометрический граф (k-NN)



После: Физически-мотивированный граф



Оценка качества

Задача модели: Бинарная классификация рёбер (принадлежит треку / не принадлежит треку)

Подход: Двухуровневая валидация напрямую оценивает полезность модели для решения физических задачи.

Метрики Машинного Обучения (на уровне рёбер)



Accuracy: Доля верно классифицированных рёбер.



Precision: Точность (доля истинных связей среди предсказанных).



Recall: Полнота (доля обнаруженных истинных связей).



F1-Score: Гармоническое среднее точности и полноты.

Физические Метрики (на уровне целых треков)



Эффективность: Доля успешно восстановленных истинных треков.



Чистота (Purity): Доля хитов от первичной частицы в восстановленном треке.



Доля ложных треков (Ghost Rate): Доля треков, не соответствующих реальной частице.

Что это даст эксперименту VM@N?



СКОРОСТЬ

Значительное ускорение реконструкции, открывающее путь к онлайн-обработке данных в реальном времени



НАДЕЖНОСТЬ

Создание независимого метода для кросс-проверки результатов, полученных классическими алгоритмами

Спасибо за внимание!