

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное
учреждение высшего образования
Национальный исследовательский ядерный университет «МИФИ»

Институт ядерной физики и технологий
кафедра №40 «Физика элементарных частиц»

ОТЧЁТ
О ПРЕДДИПЛОМНОЙ ПРАКТИКЕ

**МОДЕЛИРОВАНИЕ И РЕКОНСТРУКЦИЯ СОБЫТИЙ В МЮОННОЙ
СИСТЕМЕ УСТАНОВКИ SPD**

Научный руководитель	_____	к.ф-м.н., доц. Солдатов Е.Ю.
Научный консультант	_____	к.ф-м.н. Верхеев А.Ю.
Студент группы М23-112	_____	Осетров А.О.

Москва 2025

СОДЕРЖАНИЕ

1 ВВЕДЕНИЕ	3
2 ЭКСПЕРЕМЕНТ SPD	4
2.1 Общее описание.....	4
2.2 Мюонная система	7
3 СОЗДАНИЕ ГЕОМЕТРИЧЕСКОЙ МОДЕЛИ ПРОТОТИПА МЮОННОЙ СИСТЕМЫ SPD И МОДЕЛИРОВАНИЕ СОБЫТИЙ	10
4 АЛГОРИТМ КЛАСТАРИЗАЦИИ DBSCAN	14
5 АЛГОРИТМЫ МАШИННОГО ОБУЧЕНИЯ	17
5.1 Нейронные сети: общий принцип работы.....	17
5.2 Сверточные нейронные сети.....	19
5.3 Дерево решений.....	21
5.4 Градиентный бустинг	23
6 РАЗРАБОТКА МЕТОДА ИДЕНТИФИКАЦИИ МЮОНОВ	25
7 ЗАКЛЮЧЕНИЕ	35
СПИСОК ЛИТЕРАТУРЫ	36

ВВЕДЕНИЕ

На сегодняшний день барионная материя кажется хорошо изученной по сравнению с основными компонентами вселенной – темной материей и темной энергией. Несмотря на большие успехи, достигнутые в описании взаимодействия кварков и глюонов в рамках пертурбативного подхода КХД, вопрос о том, почему нуклоны именно такие, какими мы их видим, остается открытым. Понимание структуры и фундаментальных свойств адронов непосредственно из динамики составляющих его кварков и глюонов – одна из главных нерешенных задач КХД.

Нуклон имеет спин $\hbar/2$, что определяет его магнитный момент, а также такие фундаментальные свойства природы, как различные фазы вещества при низких температурах и стабильность известной Вселенной. Именно поэтому изучение спиновой структуры нуклона имеет особое значение. Кварковая модель успешно предсказывает большинство основных свойств адронов, таких как заряд, четность и изоспин. Однако она не позволяет объяснить спиновые свойства адронов в терминах их составных частей.

Научной установкой, которую предполагается использовать для изучения спиновой структуры нуклонов и других явлений, связанных со спином, является SPD (Spin Physics Detector) [1]. Её планируется разместить в одной из двух точек столкновения пучков коллайдера NICA, строящегося в Международной межправительственной научной организации “Объединенный институт ядерных исследований” (Дубна, Россия).

Для проектирования установки, изучения возможности решения поставленных физических задач и разработки методов регистрации и идентификации процессов в эксперименте необходимо моделирование различных детекторных систем. Работа посвящена моделированию мюонной системы SPD, с использованием программного пакета SpdRoot, и последующей разработке программного обеспечения для обработки и анализа данных эксперимента.

ЭКСПЕРИМЕНТ SPD

2.1 Общее описание

В планах коллаборации SPD установить детектор во второй точке взаимодействия строящегося коллайдера NICA (ОИЯИ, Дубна) для изучения спиновой структуры протона и дейтрона и других спиновых явлений. Планируется работать с поляризованными пучками частиц с энергией до 27 ГэВ в системе центра масс и светимостью до $10^{32} \text{ см}^{-2} \text{ с}^{-1}$.

Установка SPD, показанная на рис. 1, представляет собой универсальный 4 π -детектор с расширенными возможностями отслеживания и идентификации частиц [2].

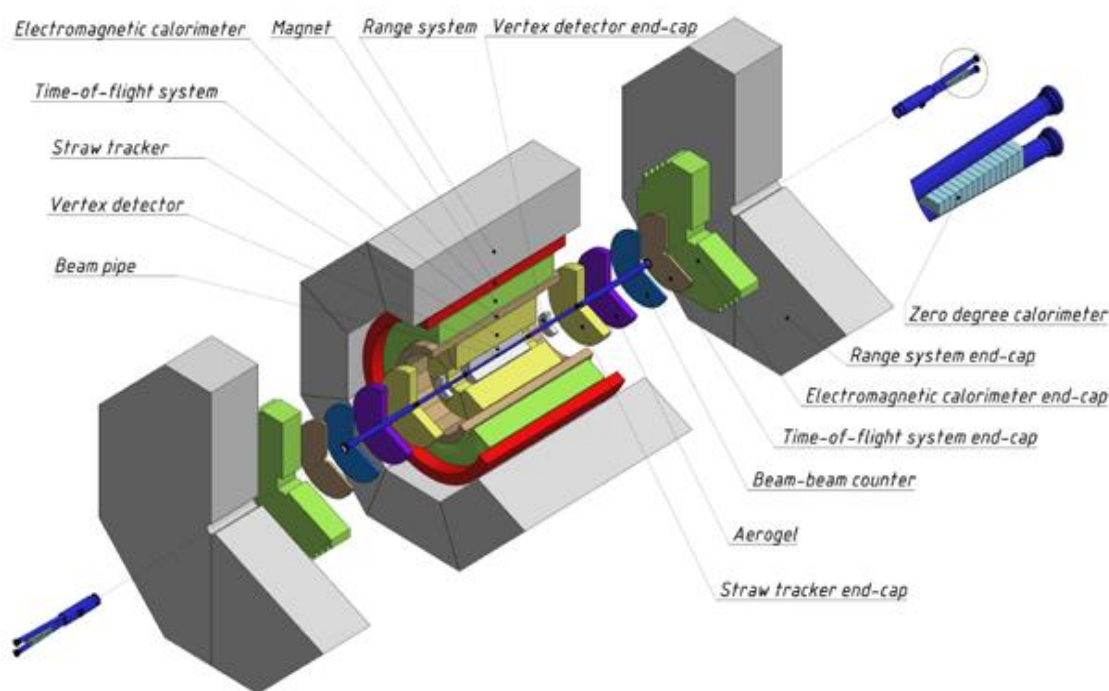


Рис. 1 Общая схема установки SPD

Прибор включает в себя следующие детекторные системы:

- магнитная система,
- вакуумная труба,
- время-пролетная система,
- вершинный детектор,
- трековый детектор,
- электромагнитный калориметр,
- счетчик столкновений пучков,
- калориметр нулевого угла,
- мюонная система.

Сверхпроводящий магнит является одной из важнейших частей установки SPD. Вместе с трековой системой он обеспечивает измерение моментов заряженных частиц с точностью около 2 %.

Вакуумная труба отделяет среду вершинного детектора от вакуума ускорителя, вследствие чего обладает большой механической прочностью и достаточно малой толщиной для минимизации многократного рассеяния и радиационных эффектов.

Цель время-пролетной системы – идентификация заряженных частиц разной массы в диапазоне импульсов до нескольких ГэВ. Планируется использовать время-пролетную систему в событиях с несколькими треками, чтобы как определить время столкновения, так и идентифицировать отдельные треки. Помимо определения типа частиц, система также сообщает время начала работы трекового детектора.

Вершинный детектор SPD – это кремниевая часть спектрометра, отвечающая за точное определение первичной точки взаимодействия и измерение вторичных вершин по распадам короткоживущих частиц (прежде всего, D-мезонов).

Целью трекового детектора является восстановление с высокой эффективностью треков первичных и вторичных частиц для измерения их импульсов по кривизне траектории в магнитном поле. Он также участвует в идентификации частиц посредством измерения энерговыведения (dE/dx).

Электромагнитный калориметр необходим для регистрации и идентификации частиц, рождающихся при столкновении адронов и вылетающих в телесный угол 4π . В его задачи входит также идентификация одиночных фотонов и нейтральных пионов, отделения фотонных ливней, порожденных π^0 -мезонами, для исключения фоновых событий.

Основными целями счетчика столкновения пучков являются локальная поляриметрия, мониторинг столкновений пучков и участие в точном определении времени столкновения, когда для этого не могут быть использованы другие детекторы (например, в случае упругого рассеяния).

Калориметр нулевого угла будет установлен в зонах разделения пучков по обе стороны от точки взаимодействия SPD. Его главные цели – измерение светимости пучка, локальная поляриметрия и создание временной метки для разделения событий.

Эксперимент SPD перекроет энергетический зазор между существующими экспериментами.

2.2 Мюонная система

Мюонная (пробежная) система SPD служит для идентификации мюонов в присутствии значительного адронного фона и оценки энергии адронов (грубая адронная калориметрия). Она также является единственным субдетектором основной части установки SPD, способным идентифицировать нейтроны путем объединения отклика с сигналами от электромагнитного калориметра и внутренними трекерами. Идентификация мюонов осуществляется посредством распознавания фрагментов трека частиц и дальнейшим их сопоставлением с треками внутри магнитов. Точное определение импульса мюона осуществляется внутренними трекерами в магнитном поле.

Одной из основных целей пробежной системы является идентификация мюонов из распадов $J/\psi \rightarrow \mu^+ \mu^-$ в присутствии фоновых пионов и продуктов их распада.

Схематичный вид системы и ее основные размеры в мм показаны на рис. 2 (а). Схема мюонной системы, состоящей из слоёв поглотителя и детекторов, показана на рис. 2 (б).

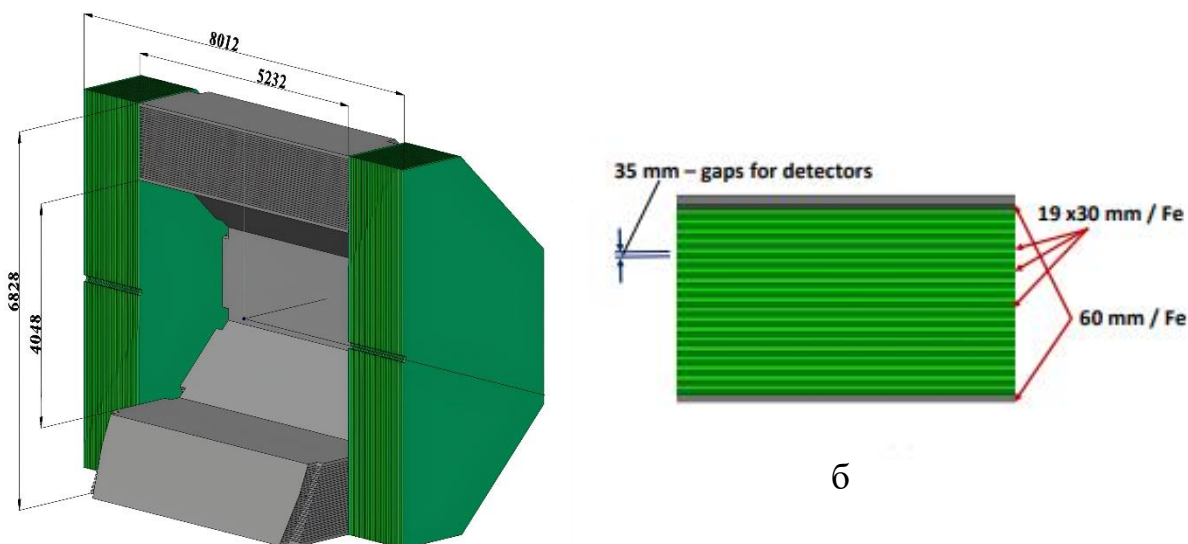


Рис. 2 (а) Схематичный вид (половина разреза) мюонной системы
Барельная часть показана серым цветом, торцевые диски – зеленым;
(б) Размеры слоев системы.

Мюонная система состоит из восьмимодульного бареля (“бочки”) и двух торцевых дисков.

Наружные слои железа толщиной 60 мм используются для скрепления модулей между собой. В межслоевые зазоры помещаются Mini Drift Tubes (MDT) детекторы и считывающая электроника. Толщина основных поглотительных железных пластин (30 мм) выбирается в сравнении с длиной пробега мюона в стали для лучшего разделения мюонов и адронов, а также для адронной калориметрии.

Общее количество MDT детекторов – около 8000 единиц. Они размещаются следующим образом: вдоль оси пучка в бареле и перпендикулярно в торцевых дисках.

Общая толщина поглотителя выбрана равной четырем длинам ядерного взаимодействия λ_I . Таким образом обеспечивается равномерная фильтрация мюонов во всех направлениях. Вместе с толщиной электромагнитного калориметра ($\sim 0,5\lambda_I$) общая толщина установки SPD получается около $4,5\lambda_I$.

Детектор MDT изначально был разработан и произведен в ОИЯИ для Мюонной системы эксперимента D0 в FNAL [3]. Позднее на его основе была создана и мюонная система для эксперимента COMPASS в CERN [4]. Модификация двухкоординатного считывания для MDT с открытой геометрией катода и внешними съемными электродами была предложена и принята коллаборацией PANDA [5] для мюонной системы их экспериментальной установки. Эта же новая версия MDT предлагается и для проекта SPD, так как имеет все необходимые характеристики – радиационную стойкость, координатное разрешение, временное разрешение, надежность, а также продвинутый уровень уже проведенных НИОКР в рамках проекта PANDA.

Сечение MDT детектора с открытой геометрией катода показано на рис. 3.

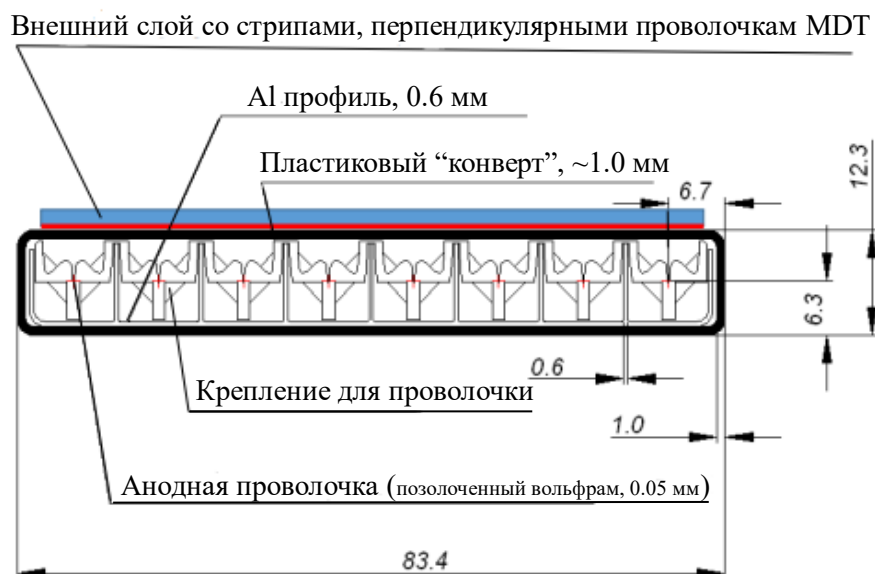


Рис. 3 Поперечное сечение MDT детектора

Детектор состоит из металлического катода (алюминиевый гребенчатый профиль из 8 ячеек), анодных проволок и пластиковой оболочки (норил) для газонепроницаемости. Гребенчатый профиль катода обеспечивает каждой анодной проволоке отдельный сигнал для получения одной из координат. Перпендикулярная стриповая плоскость используется для получения второй координаты. Форма наведенного сигнала повторяет исходный, имея противоположную полярность, но амплитуда составляет около 15% от исходного. Таким образом, требуется усиление сигнала и надлежащее экранирование внешнего электромагнитного поля.

Применение открытой геометрии катода позволяет получать вторую координату считывания. Однако это приводит к потере симметрии электрического поля в каждой из 8 ячеек детектора и, следовательно, уменьшает необходимое приложенное напряжения для того же коэффициента газового усиления по сравнению со стандартной геометрией (катод закрыт крышкой из нержавеющей стали).

СОЗДАНИЕ ГЕОМЕТРИЧЕСКОЙ МОДЕЛИ ПРОТОТИПА МЮОННОЙ СИСТЕМЫ SPD И МОДЕЛИРОВАНИЕ СОБЫТИЙ

Первый этап работы – описание прототипа мюонной системы, который в настоящее время используется для тестирования электроники, сбора космики, и интегрирование его в программную среду ROOT. Полученная геометрическая модель прототипа представлена на рис. 4.



Рис. 4 Геометрическая модель прототипа мюонной системы SPD

Сперва смоделирован MDT детектор из восьми ячеек, заполненных газовой смесью $\text{Ar} + \text{CO}_2$ (рабочее вещество газовой ячейки детектора). Затем 6 MDT объединены в одну детекторную плоскость. В железном корпусе размещено 16 детекторных плоскостей с железными поглотителями с шагом в 35 мм для электроники. В начале системы расположен так называемый *Bilayer*, состоящий из двух детекторных плоскостей, каждая из которых содержит четыре MDT детектора. За ними расположен железный лист толщиной 60 мм. Созданная модель прототипа должна стать частью разрабатываемого программного обеспечения для обработки и анализа данных проектируемого эксперимента SPD.

Далее прототип интегрирован в среду программного пакета SpdRoot с использованием методов объектно-ориентированного программирования. Необходимо было подключить все необходимые библиотеки из репозитория SpdRoot software и установить зависимости для переменных и параметров среды.

Следующим шагом произведено моделирование событий в прототипе, работу которого кратко можно описать последовательностью шагов:

- Задаются имена выходных файлов для данных и параметров,
- Создается объект SpdRunSim, который управляет симуляцией,
- Настраивается симуляция: задаются имя симулятора, используемые материалы, файл для вывода данных и конфигурация распада частиц,
- Инициализируется геометрия,
- Создаётся и настраивается генератор SpdIsotropicGenerator, который отвечает за изотропную генерацию частиц,
- Устанавливаются глобальные параметры симуляции,
- Инициализируется и запускается симуляция с заданным числом событий nEvents и отслеживанием времени выполнения,
- Параметры и данные симуляции сохраняются в выходные файлы для дальнейшей обработки и анализа.

После генерации событий запускается их визуализация для большей наглядности происходящих процессов. На рис. 5 и 6 приведены визуализации прохождения протона и мюона с энергиями 1 и 10 ГэВ через прототип. Сплошными линиями на рисунке обозначены: фиолетовая – протон, голубая – нейтрон, зеленая – электрон, желтая – позитрон, бирюзовая – мюон. Красной пунктирной линией обозначен гамма-квант, красными точками – хиты (происходящие в газовом объеме детектора). В представленных модельных событиях налетающие частицы «бьют» точно в центр детектора, тогда как в эксперименте частицы-снаряды налетают под разными углами.

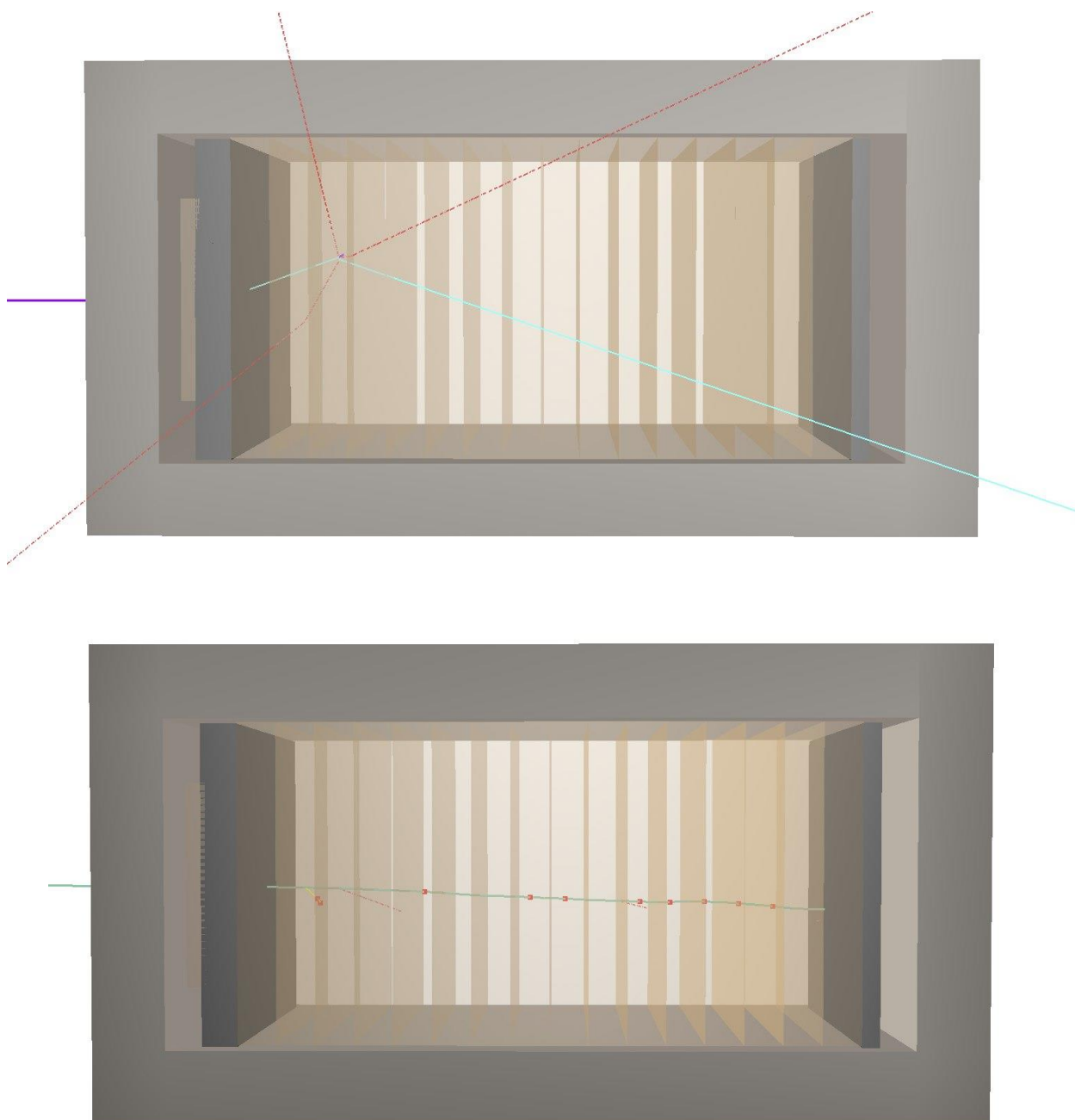


Рис. 5 Прохождения протона и мюона с энергией 1 ГэВ через прототип
Сверху – протон, снизу – мюон.

На рисунке видно, что в данном примере в результате взаимодействия протона с веществом детектора рождаются нейтроны и фотоны, покидающие пределы прототипа, в то время как при взаимодействии мюона с веществом такого не происходит – порожденные электрон и гамма-квант поглощаются внутри прототипа.

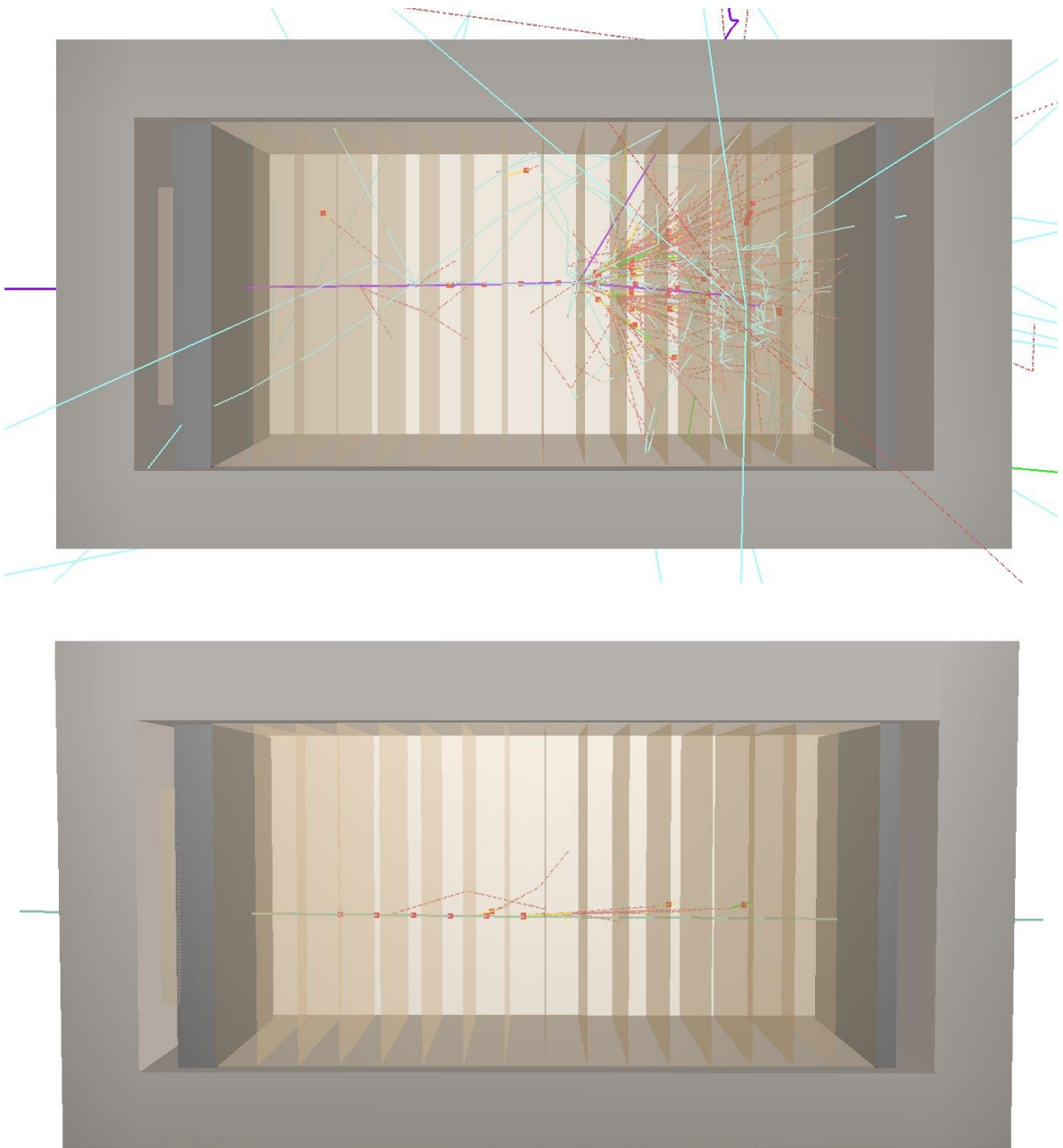


Рис. 6 Прохождения протона и мюона с энергией 10 ГэВ через прототип
Сверху – протон, снизу – мюон.

Согласно представленному рисунку, протон с энергией 10 ГэВ порождает каскад в детекторе, чего не происходит ни для протона с энергией 1 ГэВ, ни для мюона с энергией 10 ГэВ. Высокоэнергетический мюон, в свою очередь, рождает в детекторе несколько заряженных лептонов и фотонов, которые не покидают пределы прототипа. В дальнейшем полученные результаты будут сравниваться с данными, полученными при прохождении космических лучей через детектор.

АЛГОРИТМ КЛАСТЕРИЗАЦИИ DBSCAN

Для реконструкции событий по хитам в работе использовался алгоритм кластеризации DBSCAN [6]. Он относится к классу алгоритмов, которые выделяют кластеры на основе плотности точек в пространстве данных. Предположим, существует некоторая база данных D точек некоторого k -мерного пространства S . Определим несколько ключевых для DBSCAN понятий:

- ε -окрестность точки p , обозначаемая как $N_\varepsilon(p)$, определяется как $N_\varepsilon(p) = \{q \in D \mid \text{dist}(p,q) \leq \varepsilon\}$, где $\text{dist}(p,q)$ – расстояние между точками p и q .
- MinPts (Минимальное количество точек): Минимальное число точек, которое должно быть в ε -окрестности точки p , чтобы она считалась "основной точкой" (core point).
- Основная точка (Core Point): Точка, которая имеет по крайней мере MinPts соседей в пределах своей ε -окрестности.
- Пограничная точка (Border Point): Точка, которая сама по себе не является основной точкой, но находится в ε -окрестности основной точки.
- Шумовая точка (Noise Point): Точка, которая не является ни основной, ни пограничной точкой.

Работу алгоритма вкратце можно описать несколькими последовательными операциями. Сперва нужно задать входные параметры ε и minPts. В идеале соответствующие параметры должны быть оптимально подобраны для каждого кластера, но нет простого способа получить их заранее для всех кластеров базы данных точек. Однако существует простая и эффективная эвристика (представленная в разделе 4.2 статьи [6]) для определения параметров ε и minPts "самого тонкого", т.е. наименее плотного, кластера в базе данных. Поэтому DBSCAN использует глобальные значения для ε и minPts, т.е. одинаковые значения для всех кластеров.

Алгоритм запускается с произвольной точки p и проверяет, является ли она основной. Если p является основной, то все точки в её ε -окрестности добавляются в текущий кластер. Все добавленные точки (кроме p) проверяются на признаки основной. Если очередная точка является основной, то по аналогии все точки из её ε -окрестности добавляются в тот же кластер. Затем в кластер добавляются все пограничные точки. Процесс продолжается, пока все точки, достижимые из начальной основной точки, не будут добавлены в кластер.

Далее алгоритм переходит к следующей непосещенной точке и повторяет процесс до тех пор, пока все точки не будут обработаны. Шумовые точки не добавляются ни в какой из кластеров и помечаются соответствующим образом. Как уже говорилось, для DBSCAN необходимы два параметра, ε и minPts , однако эксперименты, описанные в статье [6] указывают на то, что результат работы алгоритма остается практически постоянным для разных $\text{minPts} \geq 2 * n_{\text{dim}}$, где n_{dim} – размерность пространства точек. Вследствие этого можно уменьшить количество входных параметров до одного, задав minPts равным 6 для набора точек, положение которых задается тремя координатами.

На рис. 7 приведён результат работы алгоритма на трех разных базах данных точек в сравнении с другим похожим алгоритмом – CLARANS [7].

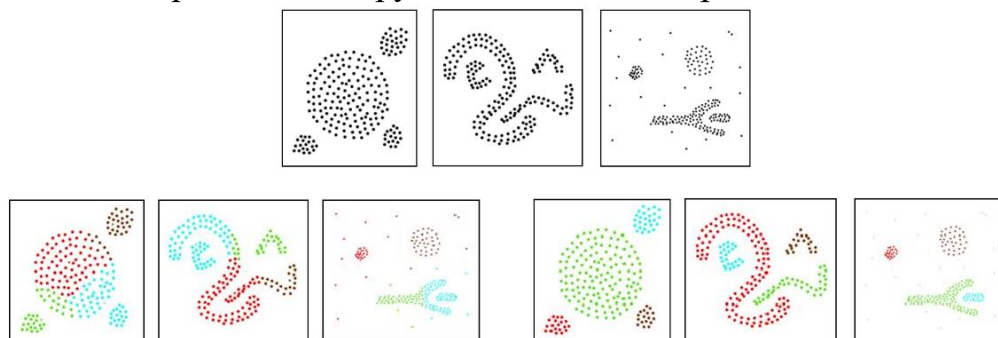


Рис. 7 Иллюстрация работы алгоритмов кластеризации
Сверху изображен оригинальный набор точек. Слева снизу – результат работы алгоритма CLARANS. Справа снизу – алгоритма DBSCAN.

Как видно из рисунка, алгоритм DBSCAN точнее группирует точки в кластеры, а также, в отличие от CLARANS, определяет шумовые точки.

Поскольку программный пакет SpdRoot разработан на языке C++, стандартные библиотеки которого не содержат встроенной реализации алгоритма DBSCAN, была разработана собственная версия данного алгоритма для последующей интеграции в SpdRoot. В целях оптимизации работы с большими объемами данных использована структура kd-дерева, реализованная в свободно распространяемой библиотеке nanoflann [8]. На рис. 8 представлен пример сравнения результата работы алгоритма DBSCAN, реализованного на C++, с результатом стандартной версии из библиотеки scikit-learn языка Python на случайном наборе точек. Аналогичные эксперименты, проведенные на множестве других случайных выборок, показали, что в каждом тесте результаты обеих реализаций полностью совпадали.

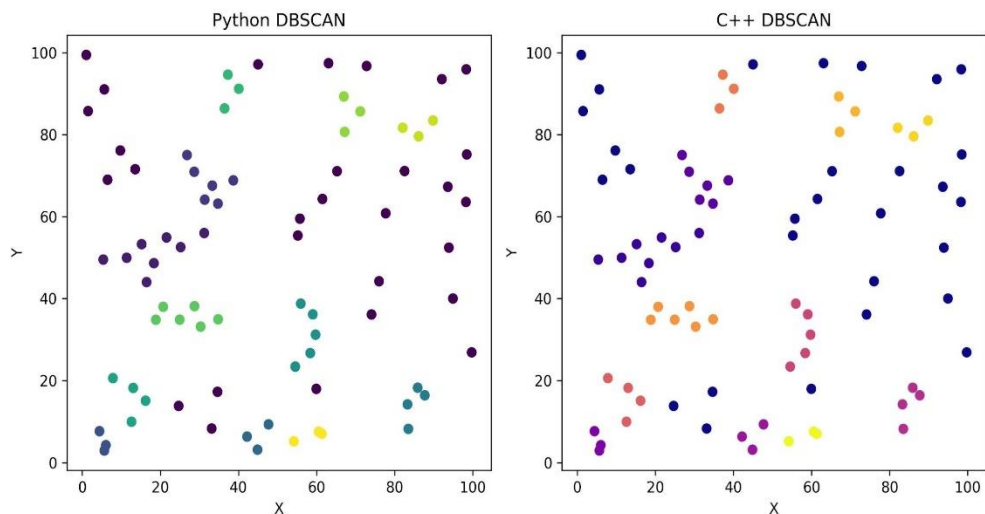


Рис. 8 Пример визуального сравнения работы двух реализаций алгоритма DBSCAN

Количество тестов n выбрано таким образом, чтобы обеспечить уровень доверия $\alpha = 0.99$ при предполагаемой вероятности ошибки в каждом тесте $p \leq 0.01$.

$$n \geq \frac{\ln(1-\alpha)}{\ln(1-p)} \approx 459 \quad (1)$$

Разработанный алгоритм интегрирован в SpdRoot для формирования кластеров хитов и решения задачи классификации частиц в мюонной системе.

АЛГОРИТМЫ МАШИННОГО ОБУЧЕНИЯ

5.1 Нейронные сети: общий принцип работы

С целью классификации треков, полученных в результате кластеризации хитов, в работе предполагается использование различных алгоритмов машинного обучения, в том числе нейронной сети.

Математическая модель нейрона, имитирующей деятельность нервных клеток, впервые была предложена Уорреном Маккаллоком и Уолтером Питтсом в работе [9]. Впоследствии Фрэнк Розенблатт разработал первую обучаемую нейронную сеть – персептрон, способную обучаться на основе примеров, что стало значимым этапом в развитии машинного обучения [10].

Искусственные нейронные сети представляют собой модели машинного обучения, основанные на принципах функционирования биологических нейронных сетей мозга. Они состоят из множества взаимосвязанных узлов — искусственных нейронов, организованных в слои. Каждый нейрон принимает входные данные, выполняет их обработку и передает результат следующим нейронам. Схема работы такого нейрона представлена на рис. 9.

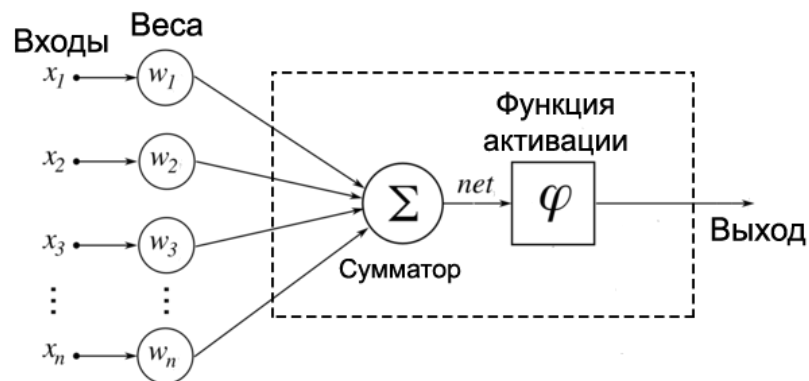


Рис. 9 Схема работы искусственного нейрона в нейронных сетях

Искусственный нейрон принимает несколько входных сигналов, которые умножаются на соответствующие весовые коэффициенты. Затем вычисляется их взвешенная сумма, к которой может добавляться смещение (bias).

Полученное значение передается на вход функции активации, определяющей выходной сигнал нейрона. Функция активации вводит в модель нелинейность, что позволяет нейросети решать более сложные задачи.

Нейроны формируют три основных типа слоев: входной, один или несколько скрытых и выходной. Исходные данные подаются на входной слой, затем проходят обработку в скрытых слоях и, в конечном итоге, поступают на выходной слой, формируя итоговый результат. Схематичное изображение структуры искусственной нейронной сети приведено на рис. 10.

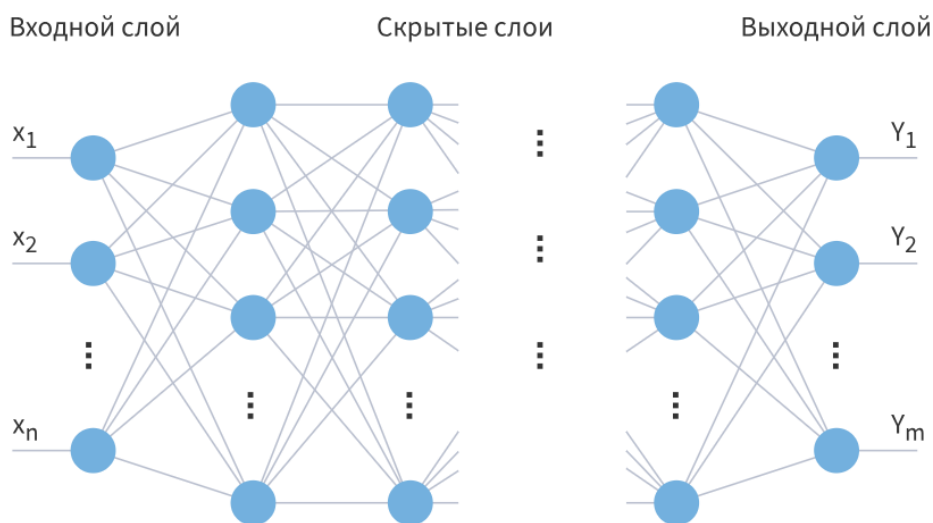


Рис. 10 Схема искусственной нейронной сети

Процесс обучения нейронной сети заключается в подборе значений весовых коэффициентов связей между нейронами с целью минимизации ошибки выходного сигнала. Одним из наиболее распространенных методов обучения является алгоритм обратного распространения ошибки [11], который корректирует веса на основе разницы между предсказанными и реальными значениями выходных данных.

Таким образом, искусственные нейронные сети моделируют процессы обработки информации, характерные для биологических систем, и находят широкое применение в задачах классификации, распознавания образов и прогнозирования.

5.2 Сверточные нейронные сети

Первым способом решения задачи классификации является применение свёрточной нейронной сети (CNN), концепция которой была впервые предложена Яном Лекуном и его коллегами [12].

Сверточные нейронные сети (Convolutional Neural Networks, CNN) представляют собой специализированную архитектуру искусственных нейронных сетей, разработанную для эффективного распознавания и обработки данных с сеточной структурой, таких как изображения. Основу CNN составляет операция свёртки, представляющая собой применение фильтров (также называемых ядрами свертки) к входным данным. Каждый фильтр представляет собой небольшую матрицу весов, которая перемещается по входному изображению, выполняя скалярное произведение между элементами фильтра и соответствующими элементами входных данных. В результате данной операции формируется карта признаков, отражающая наличие определённых характеристик в различных областях изображения. Пример работы данного механизма представлен на рис. 11.

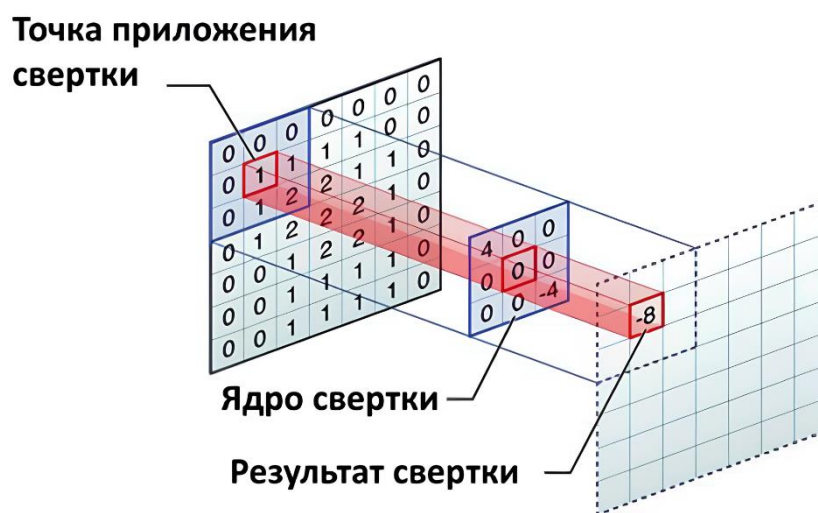


Рис. 11 Применение свертки в сверточном слое

Архитектура CNN, как правило, включает чередующиеся слои свёртки и подвыборки (пулинга), за которыми следуют один или несколько полносвязных слоёв, как показано на рис. 12. Свёрточные слои выполняют задачу извлечения локальных признаков, таких как края, углы и текстуры, в то время как слои пулинга уменьшают размер карт признаков, снижая вычислительную сложность и повышая устойчивость к смещениям и искажениям. Полносвязные слои, расположенные на завершающем этапе сети, предназначены для интеграции извлечённых признаков и выполнения окончательной классификации или регрессии.

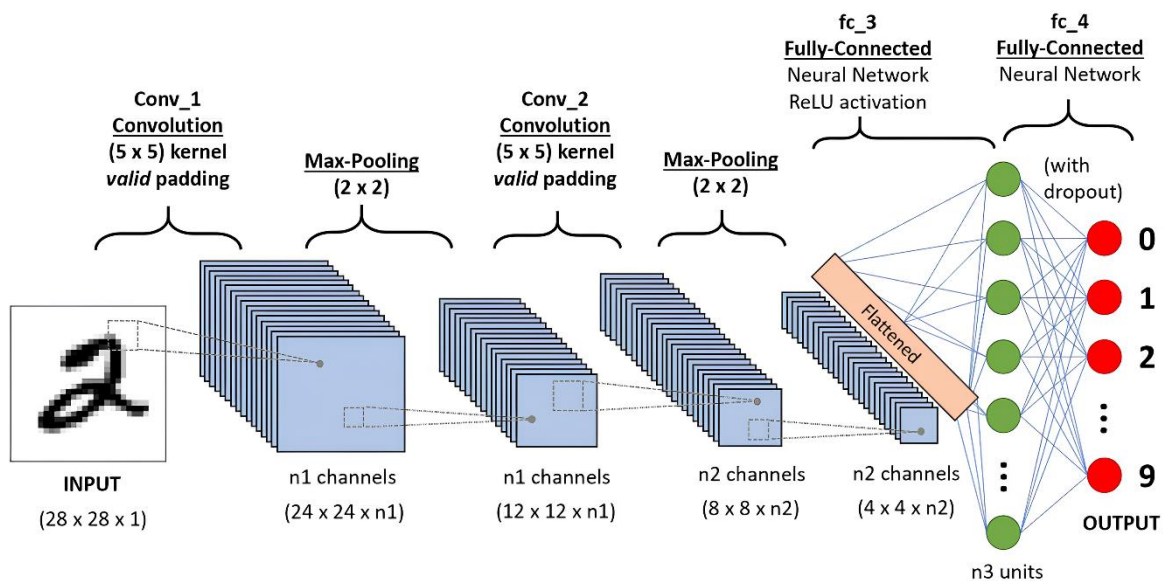


Рис. 12 Структура сверточной нейронной сети

5.3 Дерево решений

Другим алгоритмом, подходящим для решения задачи классификации, является ансамбль деревьев решений.

Деревья решений [13, 14] – это непараметрический метод обучения с учителем, применяемый для задач классификации и регрессии. Их цель – построить модель, предсказывающую целевую переменную на основе последовательности простых правил, выведенных из признаков обучающих данных. Дерево решений представляет собой иерархическую древовидную структуру, узлы которой содержат правила вида «Если ..., то ...», автоматически формирующиеся в процессе обучения. Визуализация работы данного алгоритма приведена на рис. 13.

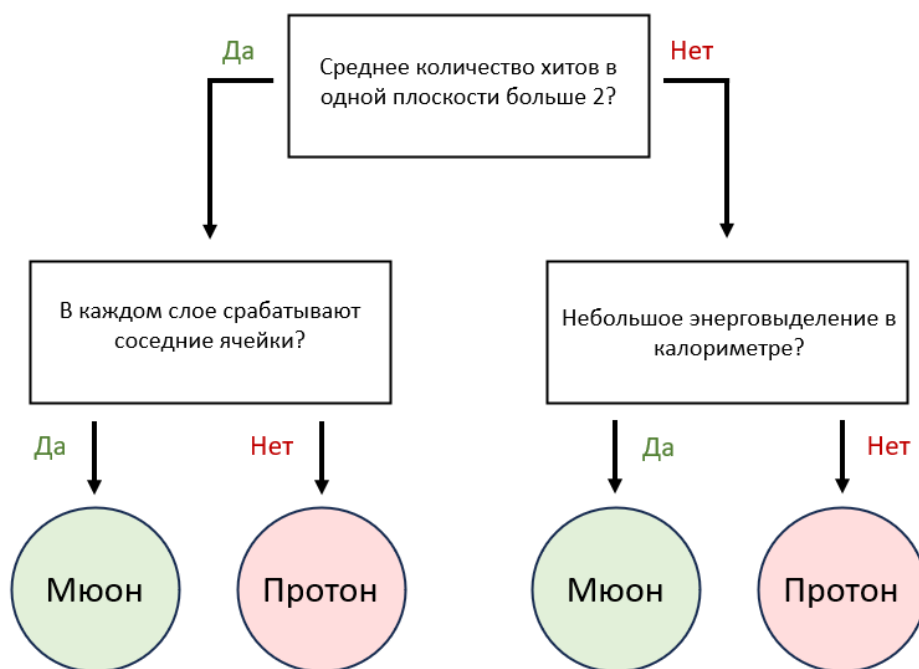


Рис. 13 Пример схемы работы дерева решений

На каждом узле дерева происходит бинарное разбиение входного множества примеров: одно подмножество содержит объекты, удовлетворяющие условию узла, другое – не удовлетворяющие. Данная процедура рекурсивно применяется к полученным подмножествам до выполнения критериев остановки алгоритма. Терминальный узел (лист) не осуществляет дальнейшего разбиения и содержит конечное решение: для задачи классификации – это соответствующий класс, для регрессии – локальное значение (или интервал) целевой переменной. Листья дерева решений обладают следующими свойствами:

1. В отличие от внутренних узлов, лист содержит не правило разбиения, а:

- Подмножество объектов, удовлетворяющих всей цепочке правил от корня до данного листа
- Финальное прогнозное значение

2. Условия попадания объекта в лист:

- Объект должен удовлетворять всем условиям на пути от корня до листа
- В силу древовидной структуры путь к каждому листу уникален
- Следствие: каждый объект может принадлежать только одному листу, что гарантирует детерминированность предсказания

Каждый новый признак в данных позволяет модели уточнить прогноз целевого значения из заданного множества классов. Полученные прогностические характеристики могут быть интегрированы в качестве входных параметров для моделей более высокого уровня. Для повышения точности предсказаний применяют ансамбли деревьев решений, где совокупный прогноз формируется на основе композиции множества отдельных деревьев.

5.4 Градиентный бустинг

Ансамблевые методы комбинируют предсказания множества базовых моделей, обученных по заданному алгоритму, для достижения лучшей обобщающей способности и устойчивости по сравнению с отдельными моделями. Классическим примером таких методов служит градиентный бустинг (Gradient Boosting) [15] – алгоритм последовательного построения линейной комбинации простых деревьев решений. На каждой итерации такой бустинг строит новое дерево, минимизирующее остаточные ошибки текущего ансамбля путем аппроксимации антиградиента функции потерь. Итоговый прогноз представляет собой линейную комбинацию предсказаний всех построенных деревьев. Ключевые преимущества алгоритма:

- Последовательная коррекция ошибок предыдущих моделей
- Возможность достижения высокой точности даже при использовании простых базовых моделей
- Естественная регуляризация за счет ограничения глубины деревьев

Хотя отдельные деревья решений легко интерпретировать благодаря возможности визуализации их структуры, модели градиентного бустинга, состоящие из сотен или тысяч деревьев, теряют это преимущество из-за своей сложности. Однако даже для таких сложных ансамблей существует возможность оценивать важность отдельных признаков, что позволяет сохранить хотя бы частичную интерпретируемость модели.

Ключевая идея заключается в том, что признаки, расположенные в верхних узлах деревьев, оказывают более существенное влияние на итоговый прогноз, поскольку они участвуют в разбиении для большей доли обучающих объектов. Напротив, признаки в глубоких узлах влияют лишь на небольшие подмножества данных и потому менее значимы.

Это наблюдение легло в основу метода MDI [14] (Mean Decrease in Impurity), который количественно оценивает важность признаков через усредненное уменьшение неопределенности (примеси) при разбиении по данному признаку во всех узлах всех деревьев ансамбля. Чем чаще признак используется для разбиения и чем выше расположены соответствующие узлы в деревьях, тем больше его итоговая важность.

Такой подход не только помогает понять относительный вклад признаков в предсказание модели, но и может быть полезен для отбора наиболее информативных признаков, что особенно ценно в задачах с высокой размерностью признакового пространства.

Градиентный бустинг, хотя и проигрывает нейронным сетям в способности выявлять сложные нелинейные закономерности и адаптироваться к различным типам данных, предлагает значительные практические преимущества при работе со структурированными данными. Главное достоинство метода заключается в его относительной простоте – для эффективной работы обычно достаточно настроить всего несколько ключевых гиперпараметров, таких как скорость обучения и глубина деревьев, в отличие от нейросетей, требующих сложной настройки архитектуры. Получаемые модели отличаются компактностью и низкими вычислительными затратами при сохранении конкурентной точности. Кроме того, градиентный бустинг предоставляет базовые средства интерпретации модели, включая оценку значимости признаков и анализ структуры отдельных деревьев, тогда как нейросети практически не поддаются содержательной интерпретации.

РАЗРАБОТКА МЕТОДА ИДЕНТИФИКАЦИИ МЮОНОВ

С целью идентификации мюонов на адронном фоне в работе применены два алгоритма машинного обучения: сверточная нейронная сеть (CNN) и градиентный бустинг (Gradient Boosting). Обучающая выборка сформирована на основе данных Монте-Карло, моделирующих прохождение протонов и мюонов с фиксированными энергиями 5 и 10 ГэВ через детектирующую систему. Проведенное сравнительное тестирование показало эффективность обоих методов, однако для окончательного выбора оптимального алгоритма требуется дополнительное исследование. В рамках дальнейшей работы планируется расширить энергетический диапазон моделируемых частиц до 1-10 ГэВ, что позволит увеличить объем и репрезентативность обучающих данных. После накопления дополнительной статистики будет выполнено повторное сравнение характеристик CNN и градиентного бустинга, что обеспечит более обоснованный выбор наиболее подходящего алгоритма для решения поставленной задачи идентификации.

Архитектура сверточной нейронной сети (CNN):

Сеть состоит из последовательности чередующихся сверточных и субдискретизирующих слоев с последующей полносвязной частью. В сверточных слоях используются ядра размером 3×3 пикселя с функцией активации ReLU, где количество фильтров последовательно увеличивается от 16 в первом слое до 32 во втором и 64 в третьем. Каждый сверточный слой сопровождается операцией пулинга (max-pooling) с окном 2×2 пикселя, что обеспечивает постепенное уменьшение пространственной размерности признаков.

После завершения сверточной части выполняется преобразование тензора в одномерный вектор (операция "вытягивания" – flatten), который подается на вход полносвязным слоям.

Полносвязная часть сети включает:

- Один скрытый слой с 128 нейронами и активацией ReLU
- Выходной слой с 2 нейронами (соответствующими классам "протон" и "мюон") с линейной активацией

Для классификации на выходе сети применяется функция softmax, преобразующая линейные выходы в вероятности принадлежности к каждому классу.

Сеть принимает на вход одноканальные изображения размером 128×128 пикселей в градациях серого. Каждое изображение представляет собой двумерную проекцию пространственного распределения хитов, зарегистрированных при прохождении частицы через детекторную систему. Примеры входных изображений приведены на рис. 14.

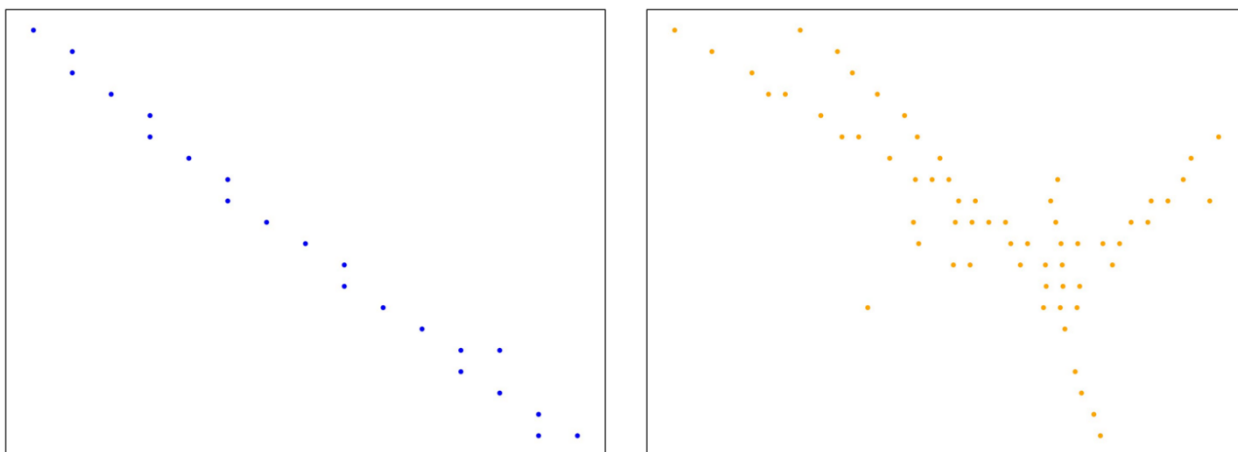


Рис. 14 Примеры входных изображений для CNN
Слева – мюон, справа – протон.

Таким образом, предложенная сверточная нейронная сеть решает задачу классификации частиц путем анализа пространственных закономерностей в распределении срабатываний детекторной системы и последующего выявления характерных признаков в картине хитов, например наличие или отсутствие каскадных процессов.

В результате тестирования алгоритма получены следующие количественные характеристики его работы:

- Точность (Accuracy): 0.969
- Полнота (Recall) для класса мюонов: 0.964
- Точность (Precision) для класса мюонов: 0.969
- F1-score для класса мюонов: 0.966

Также получены матрица ошибок (Confusion Matrix) и ROC-кривые, представленные на рис. 15 и рис. 16 соответственно.

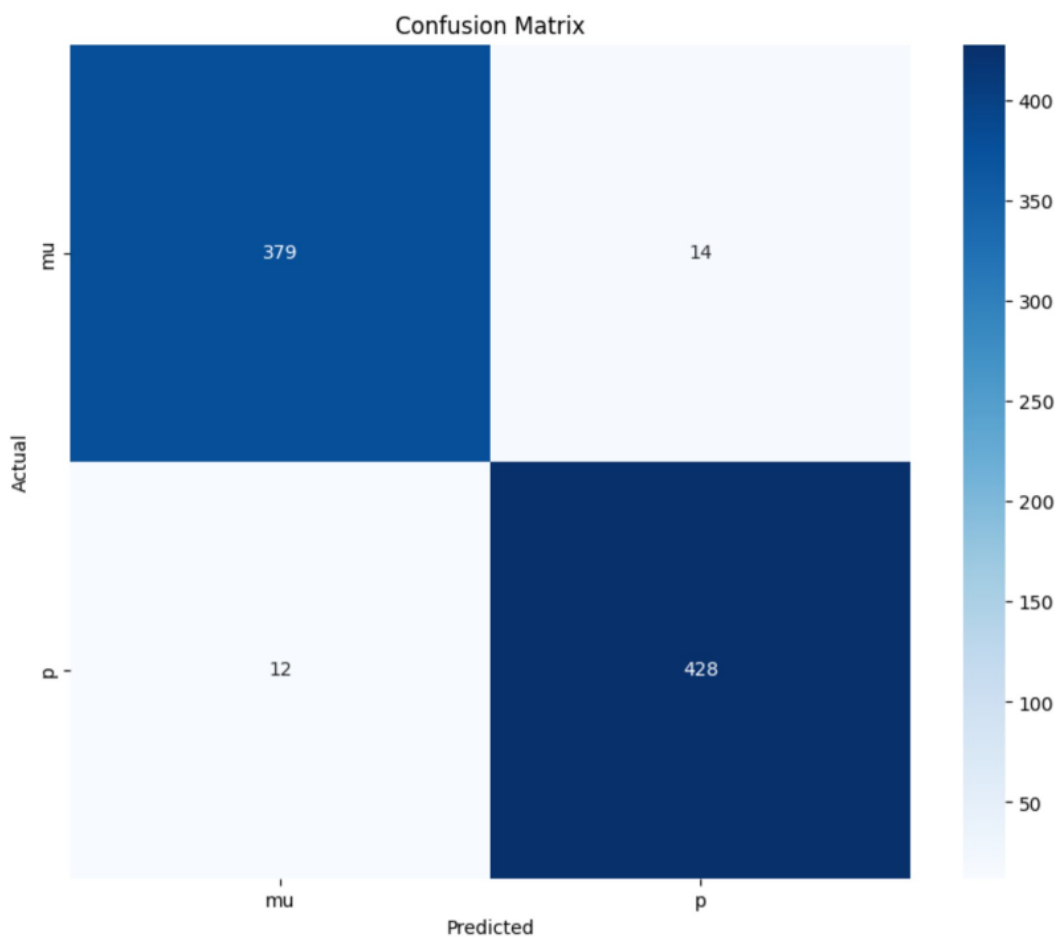


Рис. 15 Матрица ошибок для CNN

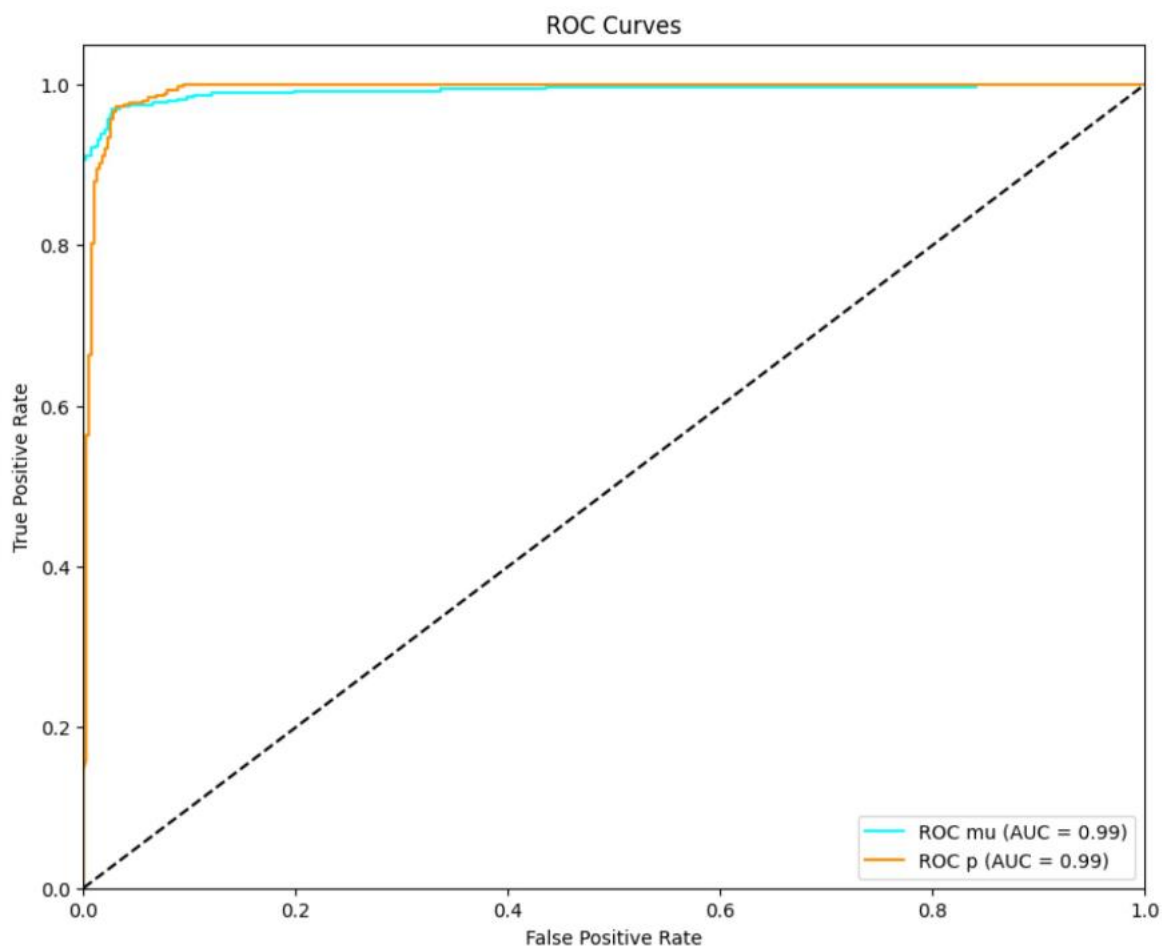


Рис. 16 ROC-кривые для CNN

В работе использовался алгоритм градиентного бустинга XGBoost со следующими настроенными гиперпараметрами:

- Количество деревьев в ансамбле: 100
- Максимальная глубина каждого дерева: 3 уровня
- Функция потерь: бинарная логистическая (binary:logistic)
- Метрика для валидации: логарифмические потери (logloss)
-

Ограничение глубины деревьев до 3 уровней обеспечивает:

- Выявление значимых признаков без переобучения
- Построение интерпретируемых правил классификации
- Эффективное вычисление предсказаний

Выбор бинарной логистической функции потерь оптимален для:

- Задачи разделения на два класса (мюоны/протоны)
- Получения вероятностных оценок принадлежности

Использование logloss в качестве метрики валидации:

- Чувствительно к качеству вероятностных предсказаний
- Жестко штрафует за уверенные ошибочные классификации

Дополнительные параметры (оставлены по умолчанию):

- Скорость обучения (learning rate): 0.3
- Доля случайных признаков: 100%
- Минимальное число образцов в листе: 1
- Коэффициент L2-регуляризации: 0

Такая конфигурация обеспечивает баланс между:

- Достаточной сложностью модели (100 деревьев)
- Устойчивостью к переобучению (ограниченная глубина)
- Интерпретируемостью результатов
- Вычислительной эффективностью

Алгоритм использовал три ключевые переменные для классификации частиц:

Интенсивность срабатываний детектора:

- Среднее количество хитов на детекторную плоскость (характеризует общую активность детектора)
- Максимальное количество хитов в одной плоскости (показывает локальные всплески активности)

Угловая характеристика трека (новый разработанный признак):

- Для каждого хита вычислялся угловой параметр в цилиндрической системе координат
- Для прямолинейного трека (характерного для мюонов) углы хитов локализуются вокруг одного значения
- Для каскадных процессов (типичных для адронов) наблюдается значительное разброс углов
- В качестве признака использовалось стандартное отклонение углового распределения σ_θ

Особенности вычисления углового признака:

Для треков в баррельной области:

- Использовалась проекция XY ($\theta = \arctan(y/x)$)
- Учитывалась только σ_{xy}

Для треков в эндкапах:

- Вычислялись обе проекции: ZX (θ_{zx}) и ZY (θ_{zy})
- Выбиралось максимальное стандартное отклонение: $\max(\sigma_{zx}, \sigma_{zy})$

Выбор проекций для анализа треков обусловлен геометрией детектора и особенностями взаимодействия частиц с веществом. Поскольку ось z детектора сонаправлена с осью пучка, в баррельной области наиболее информативной оказывается проекция XY, перпендикулярная направлению пучка. В этой проекции четко проявляются характерные особенности мюонных треков – их прямолинейность и минимальное рассеяние, а также хорошо видно развитие адронных каскадов, происходящее преимущественно в поперечной плоскости. В торцевых же частях детектора (эндкапах) из-за специфического расположения детекторных элементов – проволочных камер и стриповых детекторов – пространственная картина взаимодействий может по-

разному проявляться в проекциях ZX и ZY. Поэтому для треков в эндкапах вычисляются угловые распределения в обеих проекциях, после чего выбирается наиболее показательная из них по величине стандартного отклонения. Такой подход позволяет максимально полно учитывать особенности пространственного распределения хитов как для прямолинейных мюонных треков, так и для разветвленных адронных каскадов во всех областях детектора.

В результате тестирования алгоритма также как и для CNN получены следующие количественные характеристики работы градиентного бустинга:

- Точность (Accuracy): 0.967
- Полнота (Recall) для класса мюонов: 0.791
- Точность (Precision) для класса мюонов: 0.844
- F1-score для класса мюонов: 0.823

На основе анализа работы алгоритма построена диаграмма важности признаков (рис. 17), отражающая их относительный вклад в процесс классификации частиц.

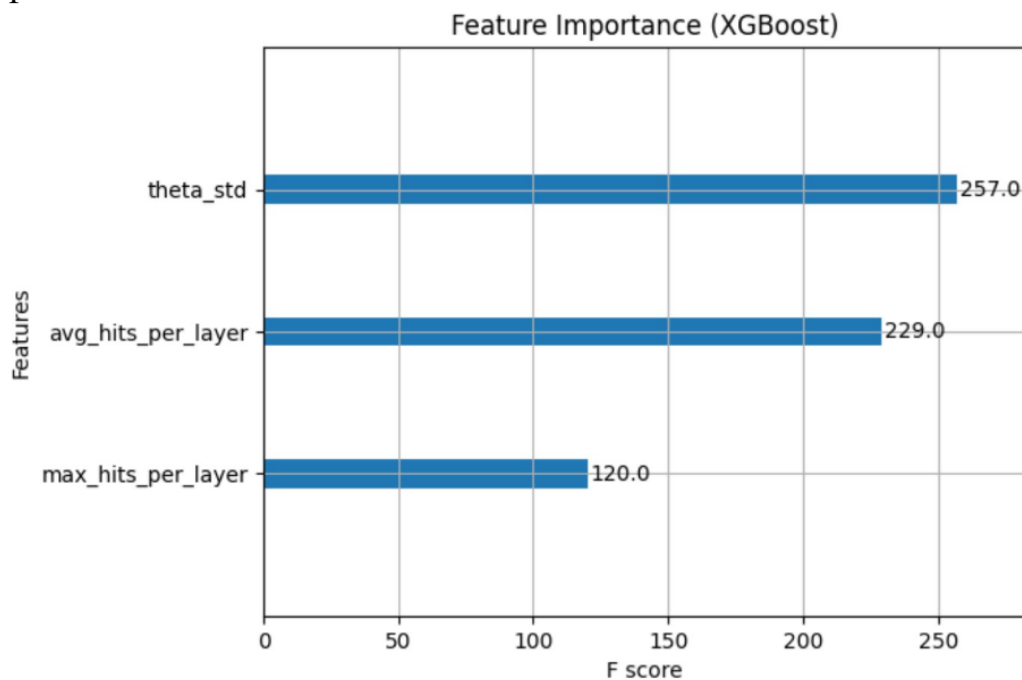


Рис. 17 Диаграмма важности признаков для модели бустинга

Как видно из диаграммы, наибольший вес имеет стандартное отклонение углового распределения хитов, что подтверждает ключевую роль пространственных характеристик трека для разделения мюонов и адронного фона. Вторым по значимости оказался признак среднего количества срабатываний в одной детекторной плоскости, который позволяет различать частицы по общей активности их взаимодействия с детектором. Максимальное количество срабатываний, хотя и уступает двум другим признакам по важности, также вносит существенный вклад в классификацию, помогая эффективно идентифицировать локальные скопления хитов, характерные для адронных ливней. Полученное распределение важности признаков хорошо согласуется с физическими ожиданиями и подтверждает обоснованность выбранного подхода к формированию признакового пространства для решения задачи идентификации мюонов.

Также получены матрица ошибок (Confusion Matrix) и ROC-кривая, представленные на рис. 18 и рис. 19 соответственно.

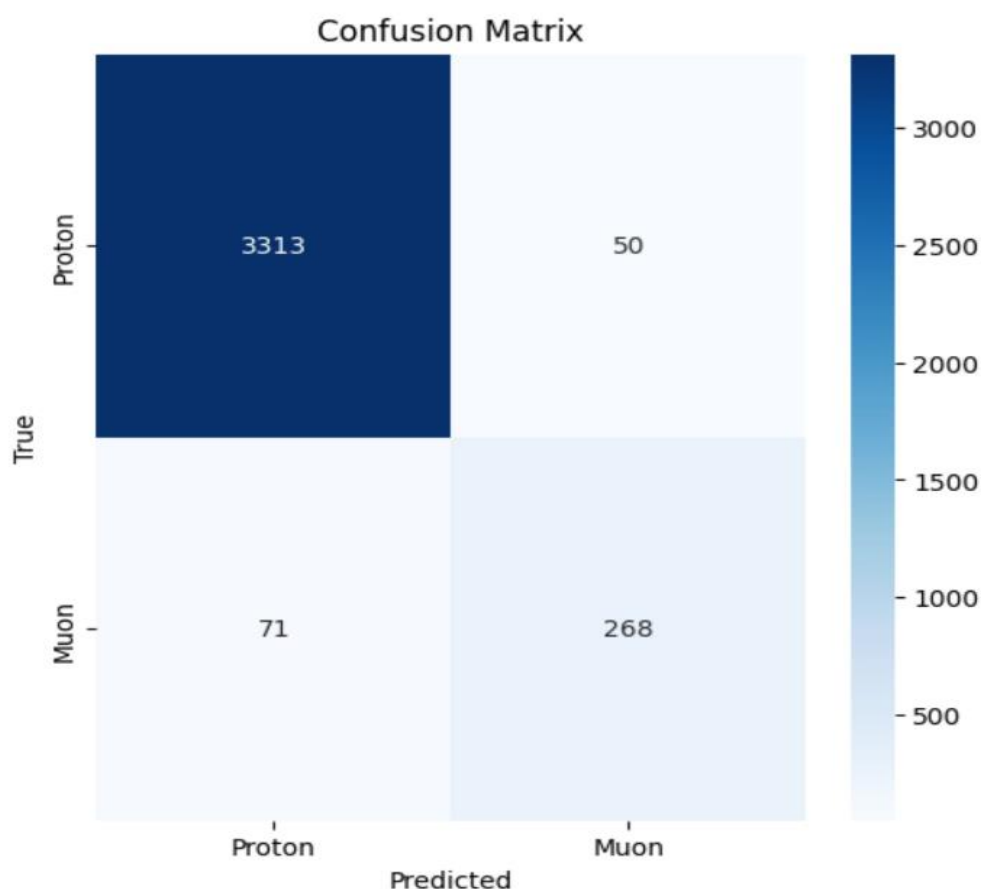


Рис. 18 Матрица ошибок для XGBoost

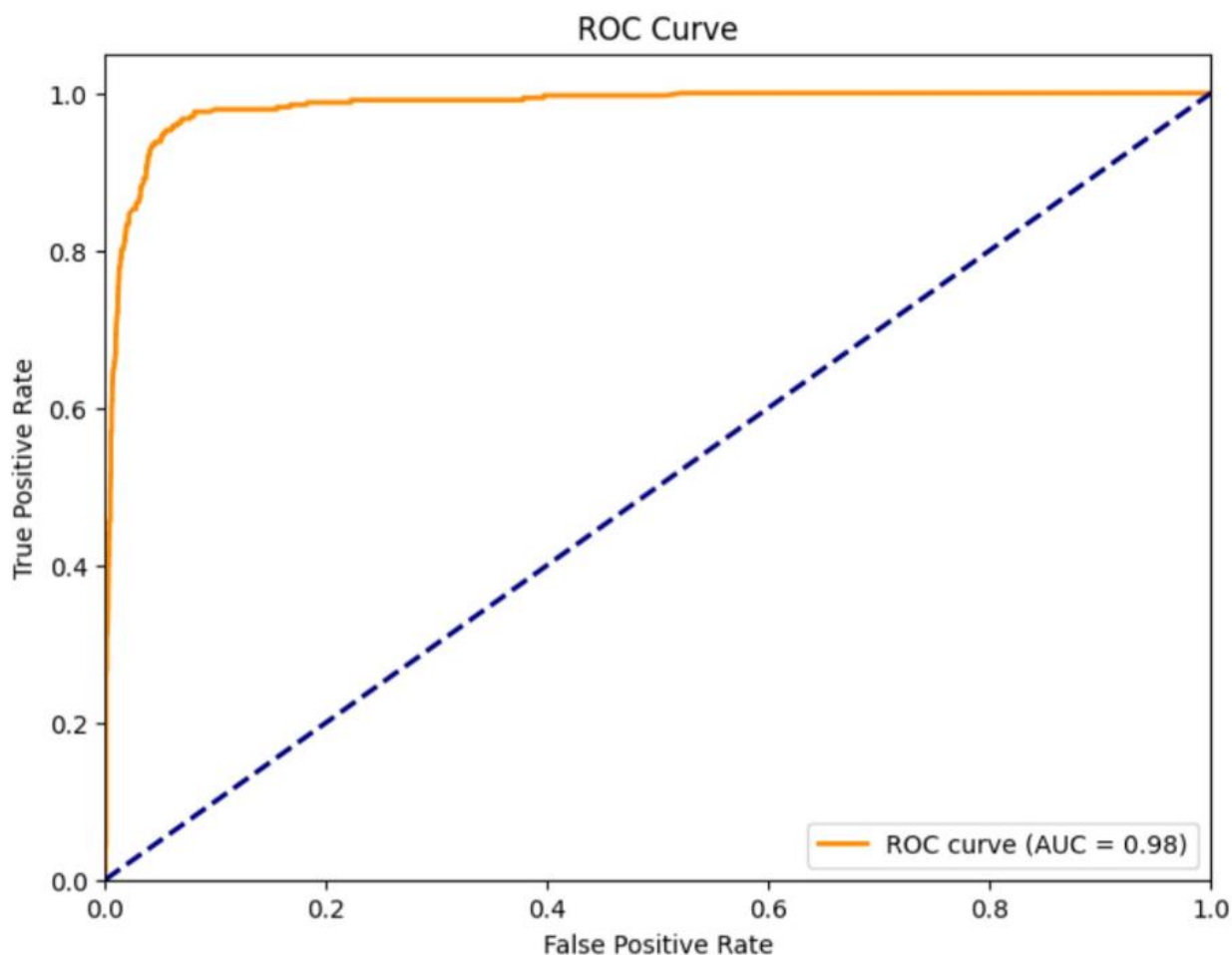


Рис. 19 ROC-кривая для XGBoost

Проведенный сравнительный анализ эффективности алгоритмов показал, что сверточная нейронная сеть (CNN) продемонстрировала более высокую точность идентификации мюонов по сравнению с градиентным бустингом. Это расхождение в производительности, вероятно, обусловлено существенным дисбалансом классов в обучающей выборке для градиентного бустинга, где количество примеров адронного фона значительно превосходило число мюонных событий.

Для устранения выявленных ограничений и повышения качества классификации предлагается следующий план дальнейших исследований:

1. Увеличение объема обучающей выборки путем проведения дополнительного моделирования мюонных событий с целью выравнивания баланса классов
2. Расширение энергетического диапазона моделируемых частиц от 1 до 10 ГэВ для повышения репрезентативности данных
3. Проведение повторного сравнительного тестирования обоих алгоритмов на обновленном наборе данных

Особое внимание следует уделить:

- Оптимизации параметров градиентного бустинга с учетом новой сбалансированной выборки и полученной диаграммы важности признаков
- Анализу зависимости качества классификации от энергии частиц
- Исследованию возможностей комбинированного подхода, сочетающего преимущества CNN и градиентного бустинга

Реализация этих пунктов позволит получить более надежные и объективные результаты сравнения алгоритмов, а также выявить оптимальный подход к решению задачи идентификации частиц в различных энергетических диапазонах.

ЗАКЛЮЧЕНИЕ

В работе произведено моделирование прототипа мюонной системы эксперимента SPD с использованием программного пакета SpdRoot. В дальнейшем геометрическую модель планируется использовать для разработки алгоритмов идентификации частиц и анализа их взаимодействия с веществом детектора. Также произведено моделирование и визуализация событий в прототипе, получен набор хитов частиц. С использованием алгоритма кластеризации DBSCAN сформированы треки для дальнейшей обработки и анализа данных моделирования. Реализация алгоритма выполнена на языке C++ и интегрирована в программный пакет SpdRoot. В ходе работы также разработаны и сравнительно проанализированы предварительные варианты алгоритмов идентификации частиц на основе CNN и градиентного бустинга. Полученные результаты, несмотря на ограниченность исходной выборки данных, демонстрируют перспективность обоих подходов.

Основным направлением дальнейших исследований станет существенное расширение и балансировка обучающей выборки путём дополнительного моделирования событий в более широком энергетическом диапазоне (1-10 ГэВ). Это позволит не только улучшить существующие модели, но и провести более обоснованное сравнение их эффективности, а также исследовать возможность создания гибридного алгоритма, сочетающего преимущества обоих методов. Особое внимание будет уделено вопросам оптимизации архитектуры нейросети и параметров бустинга, а также разработке дополнительных признаков для машинного обучения, что в конечном итоге должно привести к созданию более точного и надёжного алгоритма идентификации частиц. Полученный алгоритм также планируется интегрировать в SpdRoot.

СПИСОК ЛИТЕРАТУРЫ

- [1] Technical design report of the Spin Physics Detector / International spin physics collaboration at the collider NICA // 12 February, 2023
- [2] Conceptual design of the Spin Physics Detector/ V.M. Abazov, V. Abramov, L.G. Afanasyev // arXiv:2102.00442, January 2021
- [3] The Muon system of the run II D0 detector / V. M. Abazov et al. // Nucl. Instrum. Meth., 2005
- [4] The COMPASS experiment at CERN / P. Abbon et al. // Nucl. Instrum. Meth., 2007
- [5] Technical Design Report for the PANDA Electromagnetic Calorimeter / Rainer Novotny et al. // panda.gsi.de, August 2008
- [6] A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise / Martin Ester et al. // Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), pp. 226-231, 1996
- [7] CLARANS: A Method for Clustering Objects for Spatial Data Mining / Ng, R. T., & Han, J. // IEEE Transactions on Knowledge and Data Engineering, 14(5), 1003-1016, 1994
- [8] nanoflann: A C++ header-only fork of FLANN, a library for Nearest Neighbor (NN) with KD-trees / J. L. Blanco, P. K. Rai // Retrieved from: <https://github.com/jlblancoc/nanoflann>
- [9] A Logical Calculus of the Ideas Immanent in Nervous Activity / W. McCulloch, W. Pitts // Bulletin of Mathematical Biophysics, 1943
- [10] The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain / F. Rosenblatt // Psychological Review, 1958
- [11] Learning Representations by Back-Propagating Errors / D. Rumelhart, G. Hinton, R. Williams // Nature, 1986

- [12] Backpropagation Applied to Handwritten Zip Code Recognition / Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel // Neural Computation, 1989
- [13] Problems in the Analysis of Survey Data, and a Proposal / J. N. Morgan, J. A. Sonquist // Journal of the American Statistical Association, 58(302), 415-434, 1963
- [14] Classification and Regression Trees / L. Breiman, J. Friedman, R. Olshen, C. Stone // Wadsworth, 1984
- [15] Greedy Function Approximation: A Gradient Boosting Machine / J. H. Friedman // Annals of Statistics, 29(5), 1189-1232, October 2001