

АННОТАЦИЯ

Данная работа посвящена созданию и исследованию алгоритма обучения нейронной сети для задачи бинарной классификации частиц в рамках эксперимента SPD (Spin Physics Detector), направленного на изучение процессов физики высоких энергий. Основной целью исследования стало проектирование и реализация метода эволюционного обучения нейросетевой модели для различения пионов и мюонов на основе трековых параметров. В качестве инструмента обучения выбран генетический алгоритм, что обусловлено особенностями задачи — потенциальной недифференцируемостью функции потерь и наличием сложной структуры данных. Разработана оригинальная функция приспособленности, адаптированная к специфике задачи, а также реализованы ключевые компоненты алгоритма: турнирная селекция, адаптивные мутации, определение порогов классификации на основе средних выходов. Результаты экспериментов показали достижение точности классификации до 95% для одного из классов, несмотря на выявленные ограничения в универсальности метода. Работа демонстрирует потенциальную применимость эволюционных подходов к обучению в задачах анализа данных физики частиц.

СОДЕРЖАНИЕ

ОПРЕДЕЛЕНИЯ И ОБОЗНАЧЕНИЯ.....	5
Глава 1. Введение	6
Глава 2. Литературный обзор.....	7
2.1. Искусственные нейронные сети и их применение в физике высоких энергий	7
2.2. Ограничения градиентных методов и мотивация к использованию эволюционных алгоритмов.....	7
2.3. Эволюционные алгоритмы: общая концепция и применение к обучению нейросетей.....	7
2.4. Примеры успешного применения	8
2.5. Выводы	8
Глава 3. ПОСТАНОВКА ЗАДАЧИ И ОПИСАНИЕ ДАННЫХ.....	9
Глава 4. Методология.....	11
4.1 Выбор метода обучения	11
4.2 Архитектура нейронной сети	11
4.3 Генетический алгоритм	11
4.4 Разработка функции приспособленности	13
ГЛАВА 5. ЭКСПЕРИМЕНТЫ И РЕЗУЛЬТАТЫ	15
5.1. Подготовка данных	15
5.2. Конфигурация генетического алгоритма	15
5.3. Эксперименты с фиксированным threshold	15
5.4. Введение порога, вычисленного из данных.....	16
5.5. Замена сигмоиды на ReLU.....	16
5.6. Проверка на переобучение	17
5.7. Ограничения предложенного метода.....	18
5.8 Анализ распределения шумов на выходе нейронной сети	18
ГЛАВА 6. ЗАКЛЮЧЕНИЕ.....	20
СПИСОК ЛИТЕРАТУРЫ	21
ПРИЛОЖЕНИЕ.....	22

ОПРЕДЕЛЕНИЯ И ОБОЗНАЧЕНИЯ

SPD — Spin Physics Detector

Fitness — функция приспособленности

Threshold — пороговое значение

ReLU — Rectified Linear Unit

Dropout — метод регуляризации, который предполагает исключение случайных нейронов во время обучения сети.

Эволюционное обучение — метод оптимизации, основанный на принципах естественного отбора

Сигмоида - функция $1/(1 + e^{-x})$

Глава 1. Введение

Современные эксперименты в области физики высоких энергий сопровождаются обработкой колоссальных объемов данных. Одним из таких экспериментов является SPD, в рамках которого осуществляется анализ треков заряженных частиц, возникающих в результате столкновений. Автоматизация процесса идентификации частиц на основе характеристик их движения по трековым детекторам является актуальной задачей, способной существенно повысить эффективность исследований.

Классическим подходом в обучении нейронных сетей является градиентный спуск, однако в задачах, где целевая функция сложна, многомодальна или недифференцируема, применение этого метода затруднено. Альтернативным подходом является эволюционное обучение, имитирующее процессы естественного отбора. Эта работа посвящена созданию и тестированию алгоритма эволюционного обучения нейросети для бинарной классификации пионов и мюонов на основе их трековых параметров.

Глава 2. Литературный обзор

2.1. Искусственные нейронные сети и их применение в физике высоких энергий

Искусственные нейронные сети (ИНС) широко применяются в задачах анализа данных физики высоких энергий. Они успешно используются для идентификации частиц, восстановления треков и фильтрации событий [1][2]. В частности, эксперименты на LHC, такие как CMS, ATLAS и ALICE, активно внедряют нейросетевые методы как в оффлайн, так и в триггерных системах обработки данных.

2.2. Ограничения градиентных методов и мотивация к использованию эволюционных алгоритмов

Градиентные методы, несмотря на их эффективность в большинстве прикладных задач машинного обучения, демонстрируют ограничения в случаях с негладкими или недифференцируемыми функциями потерь, особенно типичными для физики [3]. Методы, такие как стохастический градиентный спуск и его производные, требуют наличия хорошо выраженного градиентного ландшафта, чего зачастую нет в реальных экспериментальных задачах [4].

2.3. Эволюционные алгоритмы: общая концепция и применение к обучению нейросетей

Эволюционные алгоритмы (ЭА) представляют собой мощный инструмент глобальной оптимизации, эффективно справляющийся с многомодальными функциями. ЭА использовались в задачах нейроэволюции — эволюционного обучения ИНС — начиная с 1980-х годов [5]. Среди ключевых подходов — алгоритм NEAT [6], алгоритмы коэволюции и современная работа по эволюции параметров больших сетей [7].

ЭА не требуют знания производных, а также позволяют интегрировать эвристические физические знания напрямую в функцию приспособленности, что особенно важно в научных задачах [8].

2.4. Примеры успешного применения

В работах [9][10] демонстрируется применение ЭА к задачам классификации, восстановления изображений, управления роботами и фильтрации физических сигналов. Например, в [10] применён генетический алгоритм к обучению глубоких нейросетей для анализа сигналов частиц, что позволило преодолеть ограничения традиционного подхода на нестабильных данных.

2.5. Выводы

На основе изученной литературы можно сделать вывод, что эволюционные методы обладают рядом преимуществ в контексте задач, не подходящих для градиентных подходов. Особенно это касается задач классификации и регрессии в области физики высоких энергий, где требуется высокая устойчивость к шуму, способность работать с ограниченными выборками и сложными функциями потерь.

Глава 3. ПОСТАНОВКА ЗАДАЧИ И ОПИСАНИЕ ДАННЫХ

Цель данной работы — разработка и реализация алгоритма обучения нейросети для бинарной классификации двух типов частиц — пионов и мюонов — на основе их трековых характеристик.

В качестве исходных данных использовались трековые параметры частиц. Каждая частица описывалась 12 признаками, представляющими собой числовые характеристики, отражающие форму, направление и другие параметры траектории частицы в магнитном поле детектора.

Перед началом обучения выборка была разделена на обучающую и тестовую части в соотношении 80/20. Это обеспечивало возможность оценки обобщающей способности модели на независимых данных.

Проблема бинарной классификации в данной задаче заключается в том, что распределения признаков для двух классов могут частично пересекаться. Тем не менее, предварительный анализ в виде построения гистограмм (Рис. 1) показал наличие различий в распределениях, позволяющих надеяться на успешную классификацию при условии корректного подбора метода обучения.

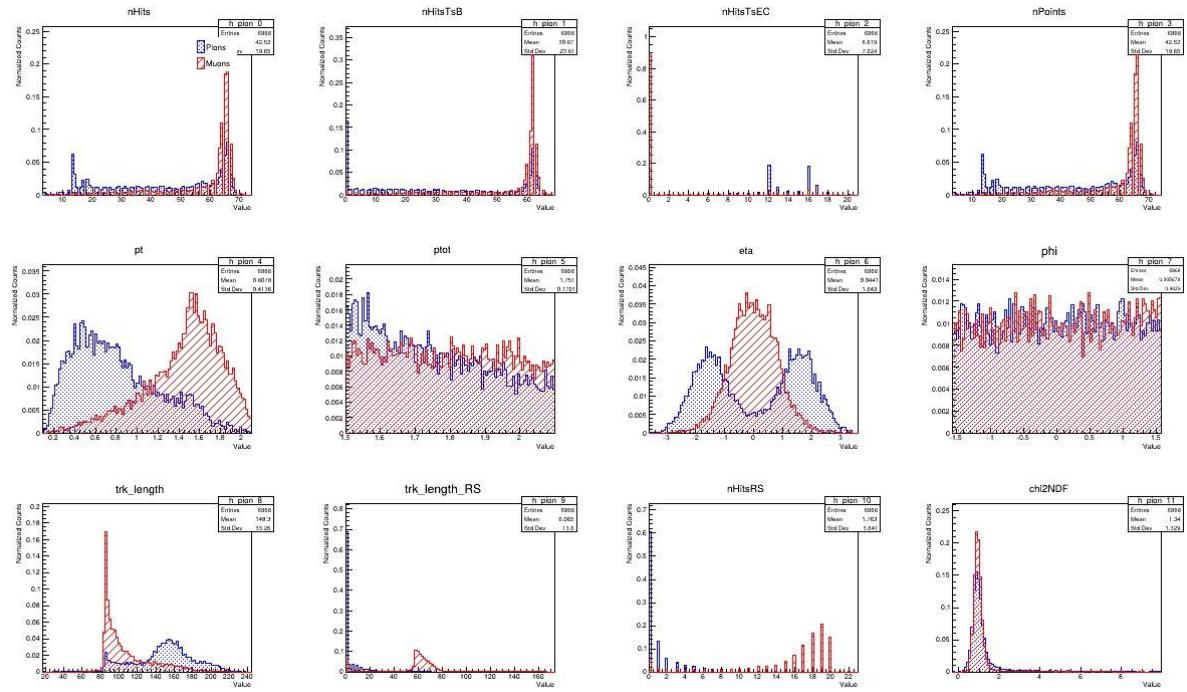


Рис. 1 Гистограммы пионов и мюонов по трековым параметрам

Таким образом, задача состоит в построении модели, которая по 12 входным признакам выдает вероятностное (или иное) значение, позволяющее отнести частицу к одному из двух классов. В отличие от стандартных подходов, данная работа делает акцент на нестандартном методе обучения — генетическом алгоритме.

Глава 4. Методология

4.1 Выбор метода обучения

Для решения поставленной задачи обучения нейросети было решено отказаться от классических градиентных методов. Причины выбора генетического алгоритма:

Возможная недифференцируемость функции потерь

Потенциально многомодальный характер пространства параметров

Устойчивость эволюционных методов к локальным минимумам

Простота внедрения эвристик и сложных критериев приспособленности

4.2 Архитектура нейронной сети

Для классификации использовалась полносвязная нейросеть с архитектурой:

Входной слой: 12 нейронов (по количеству входных признаков)

Скрытые слои: два слоя с 24 и 12 нейронами соответственно

Выходной слой: 1 нейрон

Активационные функции:

ReLU — на всех слоях, включая выходной (впоследствии заменила сигмоиду)

Сигмоида применялась в ранних экспериментах, но оказалась неэффективной из-за потери точности и сжатия диапазона выходов

4.3 Генетический алгоритм

Основные этапы работы алгоритма:

1. Инициализация популяции

Создается множество нейросетей (особей) со случайными весами.

2. Оценка приспособленности (fitness)

Каждая нейросеть тестируется на данных, и ей присваивается оценка fitness.

3. Селекция (турнирный отбор)

Выбираются "родители" для следующего поколения:

Случайно выбирается k особей из популяции (например, 5).

Среди них выбирается одна с наивысшим fitness.

Процесс повторяется, пока не отобрано нужное число родителей.

4. Скрещивание.

Гены (веса) двух родительских сетей комбинируются, создавая "потомков":

Выбирается случайная точка раздела в хромосоме (весах).

Ребенок получает часть весов от первого родителя до точки и часть от второго после

5. Мутация

Случайные изменения весов для внесения разнообразия.

6. Формирование нового поколения

Лучшие особи сохраняются (элитизм), остальные заменяются потомками.

Базовые параметры алгоритма:

Размер популяции: 100 особей

Элитизм: 10% лучших решений сохраняются без изменений

Селекция: турнирная, размер турнира 5

Мутации: аддитивные, с возможностью направленного сдвига весов

Код был реализован вручную с использованием собственных структур данных, обеспечивающих контроль за процессом отбора, мутаций и оценки приспособленности особей.

4.4 Разработка функции приспособленности

Функция приспособленности (fitness function) — ключевая часть алгоритма. Она должна поощрять точную классификацию пионов и мюонов, а также обеспечивать устойчивость к переклассификациям.

Основной подход:

Первый проход: вычисление средних выходов по классам

$$\text{Threshold} = (\text{среднее_пионов} + \text{среднее_мюонов}) / 2$$

Второй проход: начисление очков в зависимости от:

Правильности классификации

Удаленности от трешхолда

Близости к среднему значению класса

Были реализованы штрафы за неверную классификацию и бонусы за корректную, с масштабированием по расстоянию. Этот подход позволил встроить в оценку «физически осмысленную» структуру классов.

Для пионов (pions_train):

Правильная классификация ($\text{output} < \text{threshold}$):

$$\text{fitness} += 2 * (\text{threshold} - \text{output}) - |\text{output} - \text{mean_pion}|^2$$

Ложноположительная классификация ($\text{output} \geq \text{threshold}$):

$$\text{fitness} -= |\text{output} - \text{mean_pion}|^2$$

Для мюонов (muons_train):

Правильная классификация ($\text{output} \geq \text{threshold}$):

```
fitness += -(threshold - output) - |output - mean_muon|^2
```

Ложноотрицательная классификация ($output < threshold$):

```
fitness -= |output - mean_muon|^2
```

ГЛАВА 5. ЭКСПЕРИМЕНТЫ И РЕЗУЛЬТАТЫ

5.1. Подготовка данных

Для проведения экспериментов была использована выборка, содержащая параметры треков, соответствующих пионным и мюонным событиям. Каждый образец описывался 12 числовыми признаками, отражающими физические характеристики трека.

5.2. Конфигурация генетического алгоритма

В ходе экспериментов была реализована следующая базовая конфигурация генетического алгоритма:

Размер популяции: 100 особей

Доля элиты: 10%

Размер турнира для селекции: 5

Количество поколений: до 300 (в зависимости от сходимости)

Мутации: аддитивные, с изменяющейся амплитудой

Функция приспособленности рассчитывалась по специально сконструированному правилу, учитывающему разницу между выходом сети и целевыми значениями, а также отклонение от среднего выхода по классу (пояснено в главе 4).

5.3. Эксперименты с фиксированным threshold

На первом этапе были проведены эксперименты со статичным порогом классификации. Однако результаты оказались крайне неутешительными. Независимо от изменения коэффициентов функции приспособленности, архитектуры сети, количества поколений и других параметров, сеть

стабильно классифицировала все объекты как один класс (мюоны), демонстрируя точность 100% по мюонам и 0% по пионам.

5.4. Введение порога, вычисленного из данных

Для борьбы с вышеописанной проблемой была реализована идея динамического определения порогового значения. Она была основана на гистограмме (Рис. 2), построенной на откликах необученной нейросети. Кластеры накладывались друг на друга, но были отличимы.

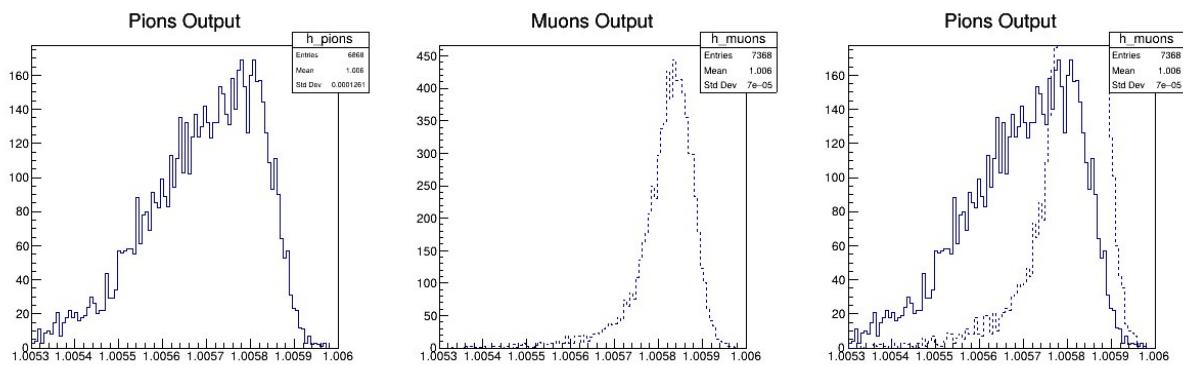


Рис. 2 Гистограммы откликов пионов и мюонов на необученной нейросети (веса выбирались случайно)

Средние выходы сети по каждому классу использовались для вычисления порога:

$$\text{threshold} = (\text{mean_output_pions} + \text{mean_output_muons}) / 2$$

Новая функция приспособленности включала штрафы и награды в зависимости от расположения выходов относительно этого порога и расстояния от средних значений по классам. Это позволило добиться частичного разделения классов.

5.5. Замена сигмоиды на ReLU

Решающим шагом стало удаление сигмоидальной функции активации и использование ReLU в качестве выходной функции. Это позволило избавиться от проблемы сжатия выходного диапазона и повысить стабильность алгоритма.

После этого архитектура сети была оптимизирована и приняла следующий вид:

Входной слой: 12 нейронов (по числу признаков)

Скрытый слой 1: 24 нейрона (ReLU)

Скрытый слой 2: 12 нейронов (ReLU)

Выходной слой: 1 нейрон (ReLU)

Такая конфигурация продемонстрировала значительно лучшие результаты:

Точность на тестовой выборке: 95% (пионы), 89% (мюоны).

5.6. Проверка на переобучение

В ходе анализа генераций не наблюдалось признаков переобучения: точность на тестовой выборке колебалась, но без явной деградации (Рис. 3). Было решено не внедрять методы регуляризации, такие как Dropout, на данном этапе.

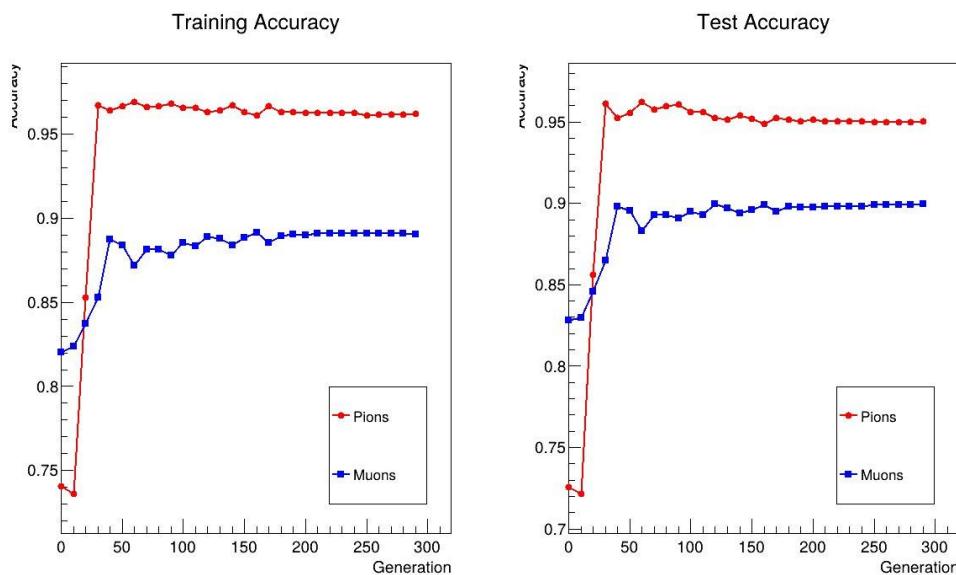


Рис. 3 Точность на обучающей и тестовой выборке в зависимости от поколения нейросети

5.7. Ограничения предложенного метода

Одним из ограничений остаётся невозможность классификации отдельных событий вне контекста выборки, поскольку определение порога завязано на распределение выходов по классам. Также в некоторых генерациях наблюдалась сильная зависимость результатов от коэффициентов функции приспособленности, что указывает на необходимость более глубокой настройки гиперпараметров.

5.8 Анализ распределения шумов на выходе нейронной сети

В ходе эксперимента были сгенерированы шумы с равномерным распределением в диапазоне от минимального до максимального значений трековых параметров пионов и мюонов. После подачи этих данных на вход обученной нейронной сети было обнаружено, что:

1. Шумы сохранили почти равномерное распределение на выходе сети, что свидетельствует об отсутствии значимой кластеризации или систематического смещения в их обработке. (Рис. 4)
2. В отличие от шумов, мюоны и пионы образовали четко выраженные кластеры, что подтверждает способность модели эффективно разделять сигналы от фона.

Полученные результаты демонстрируют, что нейронная сеть корректно обрабатывает входные данные:

1. Игнорирует равномерно распределённые шумы, не присваивая им значимых признаков классов.
2. Успешно идентифицирует реальные частицы (мюоны и пионы), группируя их в отдельные кластеры.

Это подтверждает способность модели к эффективной классификации физических событий.

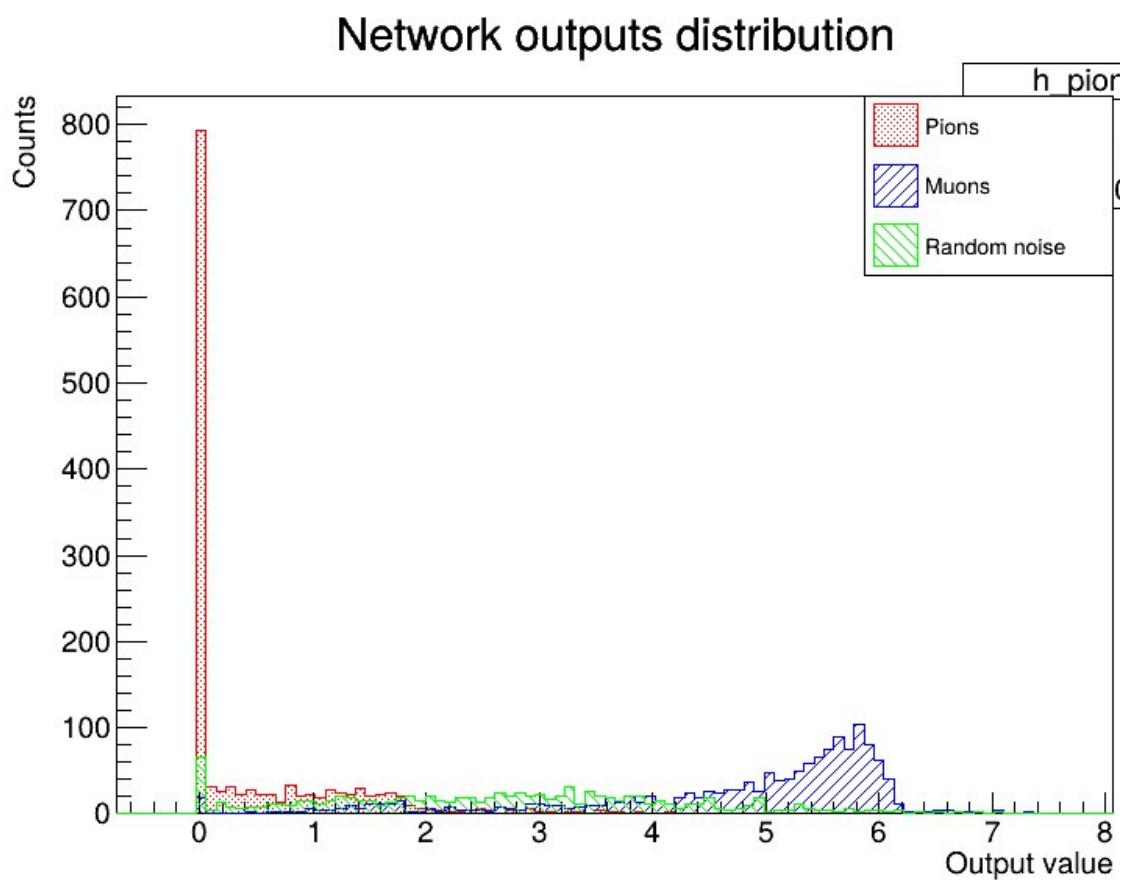


Рис. 4 Гистограмма выходов сети для пионов, мюонов и шумов для обученной нейронной сети (300 генераций)

ГЛАВА 6. ЗАКЛЮЧЕНИЕ

В рамках настоящей работы была реализована и протестирована система бинарной классификации пионов и мюонов на основе нейронной сети, обучаемой с помощью генетического алгоритма. Было показано, что стандартные методы, основанные на сигмоиде и градиентной оптимизации, оказываются неэффективны при заданной структуре данных.

Предложен и реализован подход, основанный на использовании ReLU и динамического порогового значения, вычисляемого по средним выходам классов. Этот метод продемонстрировал высокую точность и устойчивость при моделировании.

Результаты показывают, что эволюционные методы имеют высокий потенциал для задач физики высоких энергий, особенно в условиях, когда функции потерь сложны, недифференцируемы или многомодальны.

В дальнейшем планируется:

Исследовать возможность обобщения метода на многоклассовую классификацию.

Реализовать кросс-валидацию и регуляризацию.

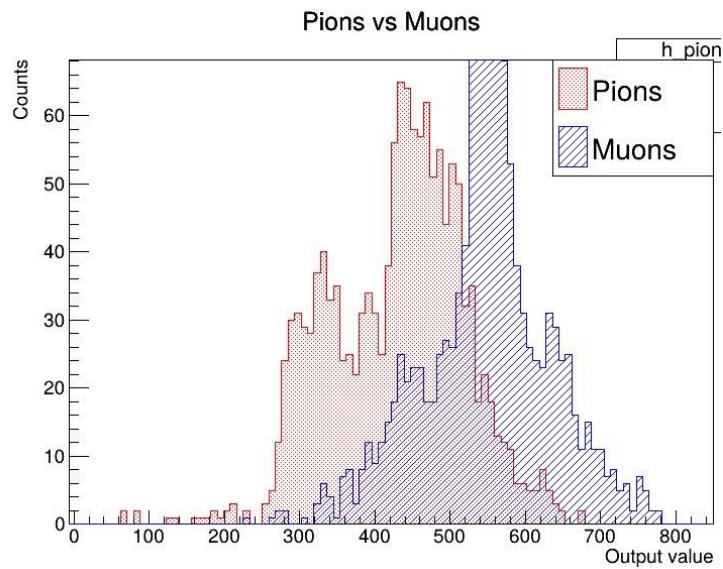
Применить алгоритм к реальным данным эксперимента SPD.

СПИСОК ЛИТЕРАТУРЫ

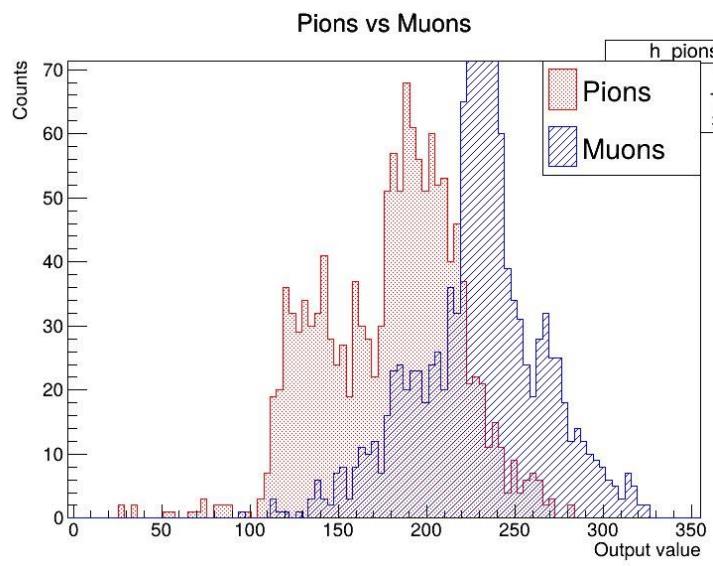
1. Goodfellow I., Bengio Y., Courville A. Deep Learning. – MIT Press, 2016.
2. Radovic A. et al. Machine learning at the energy and intensity frontiers of particle physics // Nature. – 2018. – Vol. 560. – P. 41–48.
3. Bottou L. Stochastic gradient learning in neural networks // Proceedings of Neuro-Nimes. – 1991. – Vol. 91. – P. 12.
4. Ruder S. An overview of gradient descent optimization algorithms. – arXiv:1609.04747, 2016.
5. Yao X. Evolving artificial neural networks // Proceedings of the IEEE. – 1999. – Vol. 87(9). – P. 1423–1447.
6. Stanley K. O., Miikkulainen R. Evolving neural networks through augmenting topologies // Evolutionary Computation. – 2002. – Vol. 10(2). – P. 99–127.
7. Such F. P. et al. Deep neuroevolution: Genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning // arXiv preprint arXiv:1712.06567. – 2017.
8. Floreano D., Dürr P., Mattiussi C. Neuroevolution: from architectures to learning // Evolutionary Intelligence. – 2008. – Vol. 1(1). – P. 47–62.
9. Schmidhuber J. Deep learning in neural networks: An overview // Neural Networks. – 2015. – Vol. 61. – P. 85–117.
10. T. A. Aaltonen et al. Genetic Algorithms for Track Reconstruction in High-Energy Physics // IEEE Transactions on Nuclear Science. – 2010. – Vol. 57(2). – P. 689–695.

ПРИЛОЖЕНИЕ

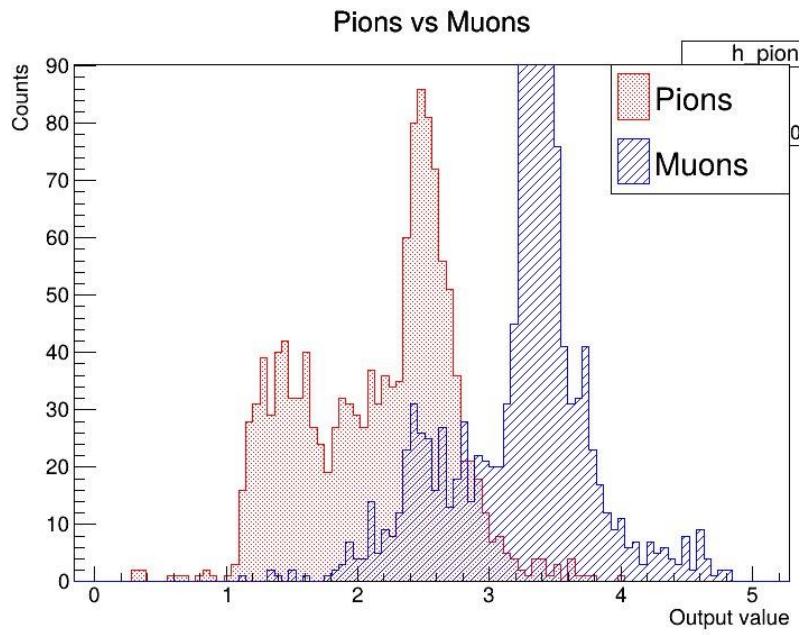
Гистограммы выходов сети для пионов и мюонов:



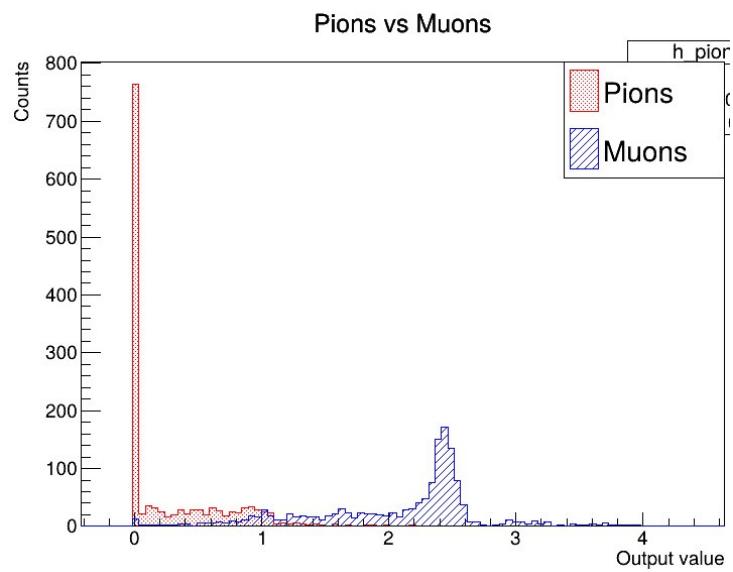
Генерация 0



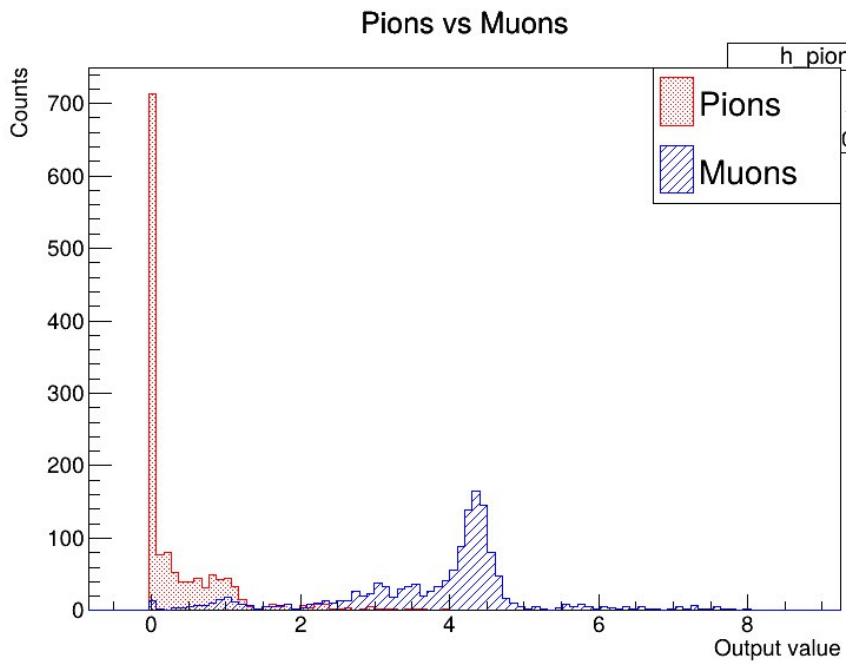
Генерация 10



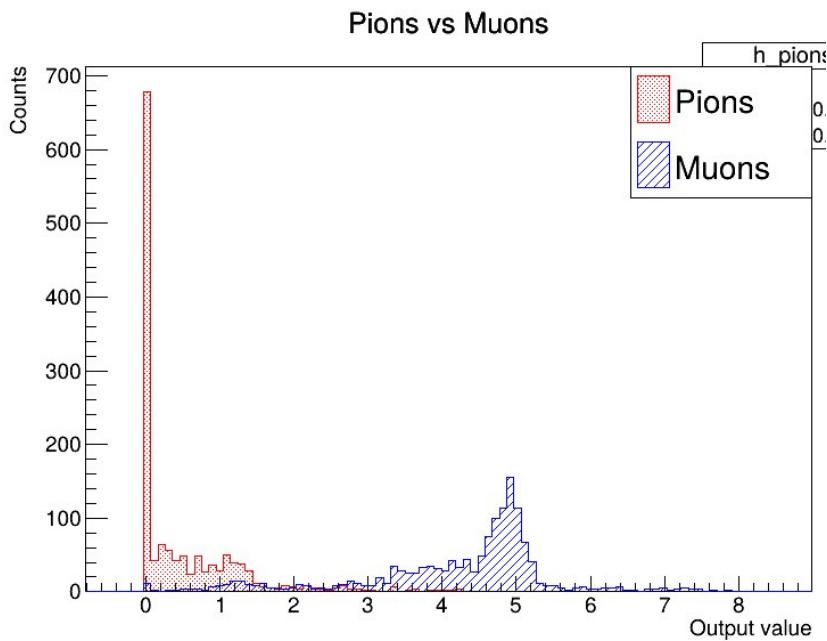
Генерация 20



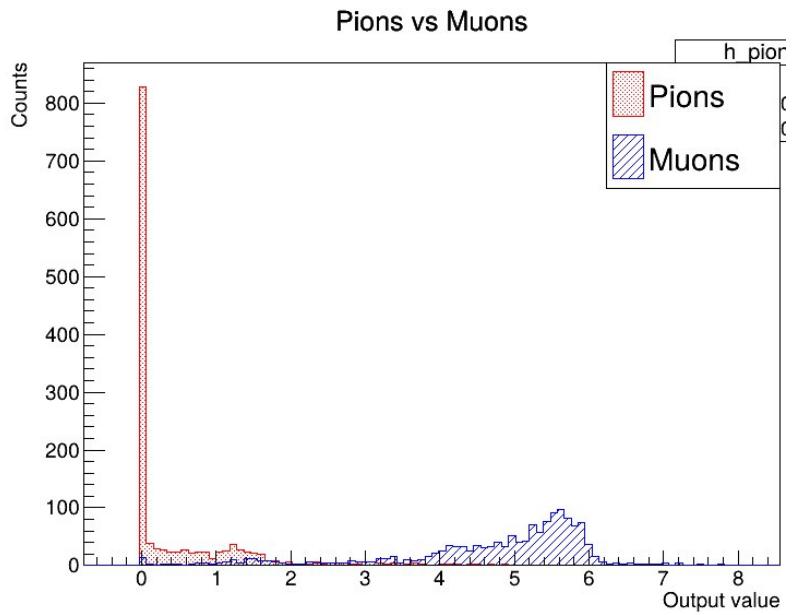
Генерация 30



Генерация 50



Генерация 100



Генерация 250