

ОБРАБОТКА И РЕКОНСТРУКЦИЯ СОБЫТИЙ НА КОЛЛАЙДЕРАХ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Пурис А.О., Голубев А.А., Дивас А.А., студенты

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Бурый А.В. – ассистент

Аннотация. В работе исследуется применение методов машинного обучения для реконструкции событий в экспериментах на коллайдерах. Рассмотрены ограничения классических алгоритмов (фильтр Калмана) в условиях высокой светимости и зашумленности. Проведен обзор современных архитектур нейронных сетей, включая графовые и сверточные сети. Показано, что методы машинного обучения позволяют эффективно обрабатывать зашумленные данные, выявлять сложные корреляции и повышать точность реконструкции, что делает их перспективным инструментом для систем сбора и обработки данных следующего поколения.

Современные эксперименты на Большом адронном коллайдере (LHC) сталкиваются с беспрецедентными объемами данных [1]. При переходе к сверхвысокой светимости (HL-LHC) частота взаимодействий достигает мегагерцового диапазона, что приводит к экспоненциальному росту комбинаторной сложности при реконструкции событий.

Классические алгоритмы, такие как фильтр Калмана, демонстрируют недостаточную эффективность в условиях pile-up (до 200 и более взаимодействий на одно пересечение сгустков), что ставит задачу поиска новых подходов [2].

Ключевой характеристикой коллайдера является энергия в системе центра масс. Для встречных пучков с энергиями E_1 и E_2 , сталкивающихся под углом θ , она вычисляется по формуле:

$$\sqrt{S} = \sqrt{2E_1E_2 + 2p_1p_2 \cos \theta} \quad (1)$$

где p_1 и p_2 — импульсы частиц.

При лобовом столкновении ($\theta = \pi$) и равных энергиях ($E_1 = E_2 = E$) выражение упрощается до $\sqrt{S} = 2E$ [3].

Величина $\sqrt{S} = 2E$ определяет порог рождения новых частиц и напрямую влияет на сложность реконструкции событий.

Основная сложность анализа заключается в том, что детекторы регистрируют не сами короткоживущие частицы, а сигналы (хиты) в трековых системах и энерговыделения в калориметрах. Для реконструкции треков заряженных частиц используется измерение радиуса кривизны R в магнитном поле B , что позволяет определить импульс частицы:

$$p = qBR, \quad (2)$$

где q — заряд, B — индукция поля, а R — радиус кривизны.

Традиционные методы, основанные на кинематической подгонке, приводят к комбинаторному взрыву и накапливают систематические ошибки в условиях высокой зашумленности [4].

В работе [5] проведен обзор методов машинного обучения, применимых для решения этих задач.

Особое внимание уделяется графовым нейронным сетям (GNN) для реконструкции треков, где данные детектора естественным образом представляются в виде графа, а также сверточным нейронным сетям (CNN) для обработки калориметрической информации как изображений [6]. Для обучения моделей используется функция потерь — кросс-энтропия, которая измеряет различие между истинным распределением вероятностей и предсказанием модели:

$$L = -\sum y_i \log(p_i) \quad (3)$$

где y_i — истинная метка класса, а p_i — вероятность, предсказанная нейронной сетью [7]. Оптимизация весов осуществляется с помощью алгоритма Adam.

В работе [8] описан процесс подготовки данных с использованием симуляций Монте-Карло (GEANT4), включающий очистку от шумов, нормализацию и разделение выборки.

Применение нейросетевых подходов позволяет не только эффективно распараллеливать вычисления, но и выявлять скрытые паттерны в данных, недоступные для аналитических алгоритмов. Установлено, что использование методов машинного обучения существенно повышает точность идентификации частиц и устойчивость реконструкции к фоновым шумам [9].

Полученные результаты подтверждают, что нейросетевые архитектуры могут стать основой для систем триггера и обработки данных следующего поколения, что особенно актуально в свете модернизации ускорительного комплекса LHC [10].

Список использованных источников:

1. Evans, L., Bryant, P. LHC Machine // *Journal of Instrumentation*. -- 2008. -- Vol. 3, No. 08. -- P. S08001.
2. Agostinelli, S. et al. (GEANT4 Collaboration). GEANT4---a simulation toolkit // *Nuclear Instruments and Methods in Physics Research Section A*. -- 2003. -- Vol. 506, No. 3. -- P. 250-303.
3. The ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data // *Journal of Instrumentation*. -- 2019. -- Vol. 14, No. 12. -- P. P12006.
4. Frühwirth, R. Application of Kalman filtering to track and vertex fitting // *Nuclear Instruments and Methods in Physics Research Section A*. -- 1987. -- Vol. 262, No. 2-3. -- P. 444-450.
5. The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker // *Journal of Instrumentation*. -- 2014. -- Vol. 9, No. 10. -- P. P10009.
6. Radovic A., Williams M., Rousseau D. et al. Machine learning at the energy and intensity frontiers of particle physics // *Nature*. -- 2018. --- Vol. 560. --- P. 41--48 .
7. Guest D., Cranmer K., Whiteson D. Deep Learning and its Application to LHC Physics // *Annual Review of Nuclear and Particle Science*. --- 2018. --- Vol. 68. --- P. 161--181.
8. CERN. GEANT4 --- Simulation Toolkit for Particle Physics Experiments. --- URL: <https://geant4.web.cern.ch/> (дата обращения: 21.02.2026).
9. ATLAS Collaboration. ATLAS Open Data Portal. --- URL: <https://opendata.atlas.cern/> (дата обращения: 21.02.2026).
10. Kingma D., Ba J. Adam: A Method for Stochastic Optimization // *arXiv:1412.6980*. --- URL: <https://arxiv.org/abs/1412.6980> (дата обращения: 21.02.2026).