

ОТЗЫВ

на диссертационную работу Гущина Михаила Ивановича
«Применение методов машинного обучения в задачах обработки и
хранения данных в экспериментах физики высоких энергий»,
представленную на соискание ученой степени кандидата
физико-математических наук по специальности
05.13.18 - Математическое моделирование, численные методы и
комплексы программ

Поиск редких процессов в ускорительных экспериментах требует высокой интенсивности столкновений ускоряемых начальных частиц, что приводит к усложнению задачи выделения треков заряженных частиц проходящих детектор. Высокая множественность вторичных частиц, рождаемых в столкновениях, приводит к необходимости развития методов выделения отдельных треков и оптимизации методов реконструкции конечных частиц. Диссертационная работа М. И. Гущина посвящена решению задачи выделения треков заряженных частиц в современных экспериментах LHCb и SHiP проводимых в ЦЕРН, Швейцария. Для решения поставленной задачи были применены передовые методы машинного обучения. Эксперимент SHiP был предложен российскими учеными [JHEP 0710 (2007) 015] и одобрен для проведения в ЦЕРН. Эксперимент LHCb является одним из четырех экспериментов проводимых на наиболее мощном современном ускорителе – Большом адронном коллайдере (БАК). Актуальность проведенных исследований и разработанных методов обусловлена их прямым применением в ведущих современных экспериментах по физике высоких энергий, и использованию наиболее эффективных и современных подходов к решению поставленных задач.

Диссертационная работа оформлена в виде введения, четырех глав, заключения и трех приложений общим объемом 260 страниц текста и 81 ссылок на литературу.

Во введении показана актуальность темы диссертации, сформулированы цели и задачи исследования, научная новизна и практическая значимость.

В первой главе рассмотрены различные методы реконструкции треков заряженных частиц в эксперименте SHiP. Приведено детальное описание методов. Сформулированы метрики сравнения различных методов. Показано применение и сравнительный анализ результатов трех методов выделения треков в эксперименте SHiP. Исследования проведены для двух обсуждаемых величин гранулярности детектора. Проведена оценка неопределенностей полученных результатов.

Вторая глава описывает оптимизацию параметров геометрии трубчатого спектрометра в эксперименте SHiP с целью улучшения качества распознавания треков и реконструкции вершин распадов частиц. Дан обзор методов оптимизации, описан процесс поиска оптимальных параметров геометрии спектрометра, приведены основные результаты тестов и сформулированы рекомендации по оптимальным параметрам детектора

Третья глава посвящена задаче идентификации типа частиц в эксперименте LHCb. В главе приведен обзор методов идентификации частиц, дано описание детектора LHCb. Представлено описание предложенных алгоритмов идентификации типов частиц и приведены результаты тестирования на смоделированных методом Монте-Карло событиях и реальных данных эксперимента. Продемонстрирована высокая эффективность предложенных методов идентификации.

В четвертой главе описано решение задачи диагностирования аномалий в системах хранения данных. Рассмотрены три класса методов диагностики аномалий, приведены результаты тестирования методов на синтетических данных и на данных реальной системы хранения для нескольких известных типов сбоев.

Основные результаты работы, выносимые на защиту, представлены в заключении. В приложениях приведены список литературы, список рисунков, таблиц, и список использованных сокращений.

В качестве замечаний к диссертационной работе следует отметить следующее. На странице 21, в формуле 1.13, нагляднее выразить определение D в виде формулы. На странице 36 нет ссылки на описание алгоритма BFGS. На странице 38 в формуле 1.43 недостаточное описание используемых обозначений. На странице 48, на рисунке 1.23 подпись оси ординат не совсем понятно отражает суть распределения, создается впечатление, что чем больше треков в событии, тем ниже относительная ошибка их восстановления. На странице 88, на рисунке 3.7 приведены только части электромагнитного калориметра детектора CMS и KTeV, хотя глава посвящена детектору LHCb. Страница 96 содержит неточную формулировку, пакет программ EvtGen моделирует распады частиц содержащих кварки тяжелых ароматов, а столкновения, адронизацию и фрагментацию кварков моделирует пакет Pythia. Страница 108 содержит неточное описание пакета Keras и не приведена ссылка на описание этого пакета. В соответствии с документацией, программный пакет Keras предоставляет интерфейс к реализациям нейронных сетей в пакетах TensorFlow, Theano, а не реализацию нейронных сетей.

Однако, высказанные замечания не влияют на положительную оценку диссертационной работы. Диссертация представляет собой законченную научно-исследовательскую работу. Результаты, полученные в ходе данного исследования, являются новыми, использованные методики обеспечивают их достоверность. Положения, выносимые на защиту, основаны на результатах, полученных с помощью современных и эффективных методов машинного обучения, а также с помощью стандартных программных пакетов реализующих эти методы, в силу чего являются обоснованными и достоверными. Результаты работы имеют существенное значение как для экспериментальной физики элементарных частиц, так и для развития методов машинного обучения и могут быть использованы в последующих экспериментах и различных прикладных областях применения методов машинного обучения.

Текст автореферата в полной мере отражает содержание диссертации, сформулирован личный вклад автора, материал диссертации отражен в публикациях указанных в автореферате.

Диссертация Гущина М. И. на соискание ученой степени кандидата наук является научно-квалификационной работой, в которой содержится решение ряда задач развития и практического применения методов машинного обучения при проведении современных коллайдерных экспериментов. Диссертация представляет собой законченную квалификационную работу, имеет научную новизну и практическую значимость, соответствует критериям установленным Правительством Российской Федерации, а ее автор Гущин Михаил Иванович заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 05.13.18 Математическое моделирование, численные методы и комплексы программ.

Заведующий Лабораторией электрослабых
и новых взаимодействий ОЭФВЭ НИИЯФ МГУ,
к.ф.-м.н.



Л.В. Дудко

Подпись Л.В. Дудко заверяю
Секретарь Ученого совета НИИЯФ и ОЯФ"
к.ф.-м.п.



Е.А. Сигаева

119991, ГСП-1, Москва, Ленинские горы, дом 1, строение 2
Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования Московский государственный университет имени М.В. Ломоносова, Научно-исследовательский институт ядерной физики имени Д.В. Скобельцына (НИИЯФ МГУ), <http://www.sinp.msu.ru>, info@sinp.msu.ru
Лев Владимирович Дудко
Тел.: +7 495 939 58 81
E-mail: lev.dudko@cern.ch