

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

УДК 53.05, 53.07

ОТЧЕТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

**Создание нейронной сети для классификации
лептонных распадов W бозона в протон-протонных
столкновениях эксперимента ATLAS**

Научный руководитель

_____ Д. Е. Пономаренко

Выполнил

_____ Г. А. Толкачёв

Москва 2020

Оглавление

1	Цель работы	2
2	Введение	3
2.1	Стандартная модель	3
2.2	Лептонная универсальность	4
2.3	Машинное обучение	5
3	Детектор ATLAS	10
4	Использованные данные	10
4.1	Экспериментальные данные	11
4.2	Монте–Карло моделирование	11
4.3	Предварительный отбор	13
4.4	Кинематические распределения	14
5	Обучение нейронной сети	16
6	Оценка отклика нейронной сети	19
7	Заключение	22
	Список использованных источников	23

1 Цель работы

Целью данной работы является создание нейронной сети, на основе которой будет осуществляться классификация событий лептонных распадов W бозона в протон–протонных столкновениях эксперимента ATLAS, а также оценка отклика нейронной сети при классификации с учетом коэффициентов коррекции и без учета. Кроме того, целью является получение начальных сведений о нейронных сетях и ознакомление с объектно ориентированным программированием в языке Python.

2 Введение

2.1 Стандартная модель

Исследования в физике элементарных частиц привели к созданию теории взаимодействия частиц на субъядерном уровне, которую принято называть Стандартной моделью (рис. 1). Стандартная модель позволяет теоретически предсказать свойства различных процессов в физике элементарных частиц. В рамках Стандартной модели имеется 2 типа элементарных частиц: бозоны и фермионы. Фермионы имеют полуцелый спин. Сами фермионы делятся на две подгруппы: кварки и лептоны. Лептоны делятся на 2 типа: электрически заряженные частицы (e, μ, τ) нейтральные частицы — нейтрино (ν_e, ν_μ, ν_τ). Кварки являются массивными частицами, имеющие электрический и цветовой заряд, всего их 6 (u, d, c, s, t, b). Все фермион по возрастанию массы делятся на три поколения.

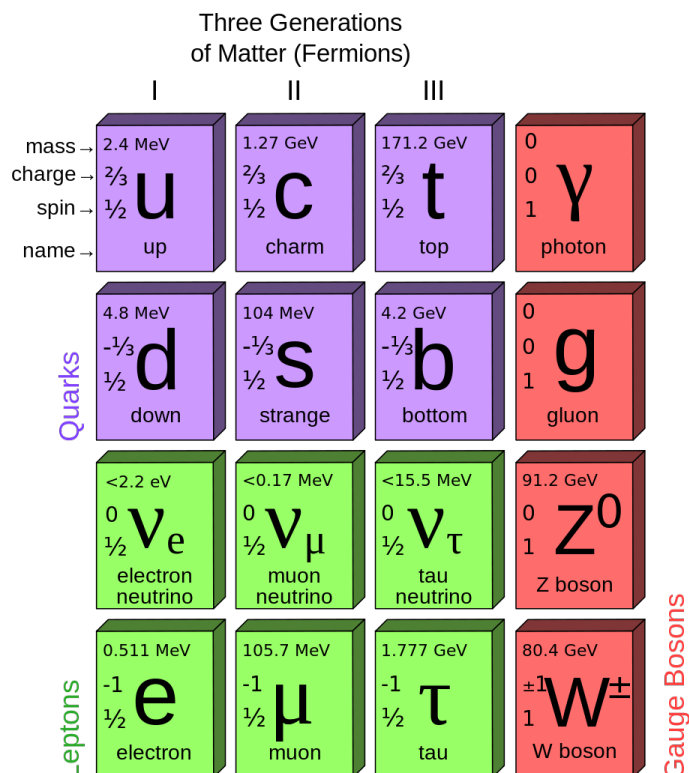


Рисунок 1 – Иллюстрация стандартной модели

В стандартную модель входит 3 вида взаимодействия: электромагнитное, слабое и сильное. В электромагнитном взаимодействии участвует ча-

стицы, имеющие электрический заряд. Переносчиком электромагнитного взаимодействия является фотон. Слабое взаимодействие связано со слабым изоспином. Переносчиком слабого взаимодействия являются Z^0 и W^\pm бозоны. Электромагнитное и слабое взаимодействия являются составными частями электро-слабого взаимодействия. Сильное взаимодействие является короткодействующим взаимодействием, благодаря которому кварки являются связанными внутри нуклонов и других адронов. Сила этого взаимодействия намного превосходит силу электромагнитного и слабого взаимодействия. Из фундаментальных частиц в сильном взаимодействии участвуют только кварки и глюон. Глюон является переносчиком сильного взаимодействия. Кварки реализуют сильное взаимодействие, обмениваясь глюонами. Это взаимодействие обусловлено наличием у кварков характеристики, называемой цветовым зарядом. Кварк может иметь один из трех цветовых зарядов. Сильное взаимодействие — это цветное взаимодействие, в процессе которого кварки обмениваются цветом, переносимым глюоном. Поэтому глюон также несет цветовой заряд. Существует 8 вариантов цветовых зарядов глюонов. В рамках Стандартной модели переносчики трех фундаментальных взаимодействий: электромагнитного, слабого и сильного являются калибровочными бозонами. Но, не смотря на огромный успех, как в теоретической части, так и в экспериментальной, стандартная модель не является окончательной теорией всего. Например, в СМ не входит гравитационное взаимодействие, а так же темная материя и энергия.

2.2 Лептонная универсальность

Три поколения упорядочены по массе заряженного лептона в диапазоне от 0.511 МэВ для e до 105 МэВ для μ , и 1.777 для τ [6]. Различия в массе приводят к совершенно разным временам жизни. От стабильного e до 2.2 мкс для μ , и 0.29 пс для τ . Лептоны участвуют в электромагнитных и слабых, но не сильных взаимодействиях, тогда как нейтрино участвуют только в сильном взаимодействии. Стандартная модель предполагает, что эти взаимодействия заряженных и нейтральных лептонов универсальны, т.е. одинаковы для трех поколений. Количество поколений лептонов пока

не объяснено в рамках стандартной модели. Почти все наблюдаемые во Вселенной процессы выглядели бы точно так же, если бы существовало только одно поколение лептонов. Имеются указания на отклонения [3] от лептонной универсальности в отношениях сечения лептонного распада W в τ к полному сечению лептонного распада W бозона (рис. 2).

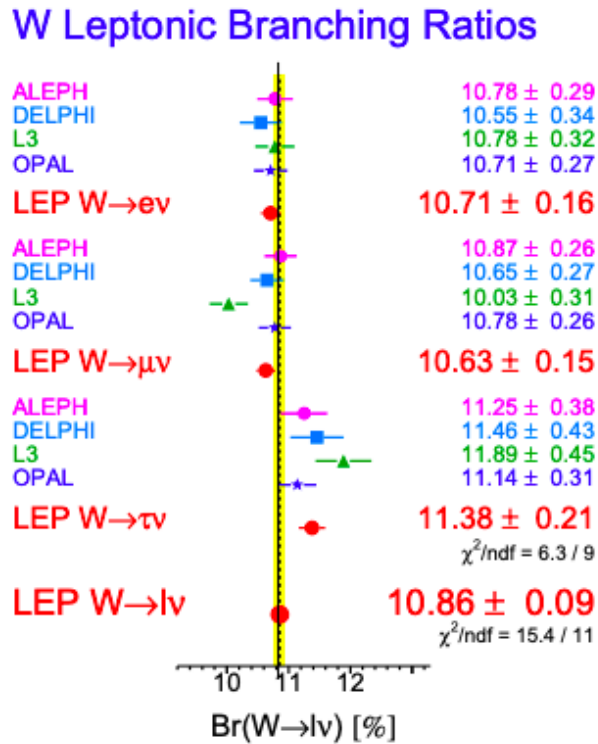


Рисунок 2 – Отношение сечения определенного канала распада к полному сечению распада W бозона

2.3 Машинное обучение

Роль машинного обучения в современном мире трудно переоценить, с помощью него можно выполнять различные задачи, как в физике высоких энергий, так и в других областях науки. В данной работе использовалась TMVA (Toolkit for Multivariate Data Analysis with ROOT) [8] — open-source библиотека алгоритмов машинного обучения, которая идёт в дополнение к пакету анализа больших данных ROOT. На ее основе осуществлялась классификация событий в лептонных распадах W бозона. На этапе обучения, нейронной сети были поданы размеченные данные. После завершения обучения нейронная сеть получала данные без меток. Задачей нейронной сети

было выполнить классификацию тестовой выборки на основе обучения, которое было совершено ранее.

Искусственная нейронная сеть

Искусственная нейронная сеть — это последовательность нейронов, соединенных между собой синапсами. Синапсом является соединение, которое используется для отправки или получения информации между нейронами. Нейрон является базовым компонентом нейронной сети. Он принимает входные данные, производит с ними определенные математические операции, затем выводит результат или передает по синапсу к следующему нейрону. У нейрона есть n входов x (рис. 3), у каждого из которого

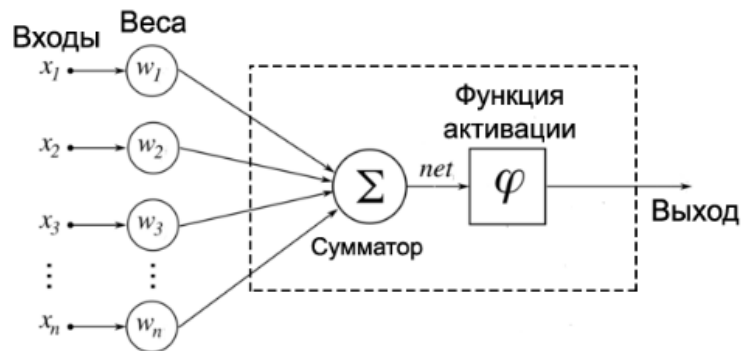


Рисунок 3 – Модель искусственного нейрона

есть вес w_n , на который умножается сигнал, проходящий по связи. После этого взвешенные сигналы $x_n w_n$ направляются в сумматор, который агрегирует все сигналы во взвешенную сумму. Результатом работы сумматора является число, называемое взвешенной суммой.

$$x_1 w_1 + x_2 w_2 + \dots + = \sum_{i=1}^n x_i w_i$$

Взвешенная сумма (net) - сумма входных сигналов, умноженных на соответствующие им веса.

$$net = \sum_{i=1}^n x_i w_i$$

Для формирования выходного сигнала используют функцию активации $\varphi(net)$, которая преобразует взвешенную сумму в какое-то число, которое

и будет являться выходом нейрона. а

$$out = \varphi(net)$$

Для разных типов нейронов используют самые разные функции активации, но одними из самых популярных являются:

- Сигмоидальная функция
- Функция единичного скачка
- Гиперболический тангенс

Персептрон

Персептрон является простейшим видом искусственной нейронной сети. Многослойный персептрон - это класс искусственных нейронных сетей прямого распространения, состоящих как минимум из трех слоёв: входного, скрытого и выходного. За исключением входных, все нейроны используют нелинейную функцию активации. При обучении MLP используется обучение с учителем и алгоритм обратного распространения ошибки. В качестве активационных функций нейронов используются: сигмоидальная функция или гиперболический тангенс. В этой работе использовался персептрон с одним скрытым слоем, количество нейронов скрытого слоя равно $k + 1$, где k количество переменных, использованных при обучении. В качестве функции активации был использован гиперболический тангенс, так как он обеспечивает более быстрое обучение.

Функция потерь

Процесс обучения нейронной сети заключается в подстройке ее внутренних параметров под конкретную задачу. Алгоритм работы нейронной сети является итеративным, его шаги называют эпохами. Эпоха - одна итерация в процессе обучения, включающая предъявление всех примеров из обучающего множества и, возможно, проверку качества обучения на кон-

трольном множестве. Выходной сигнал нейронной сети определяется как

$$y_{out} = \sum_{j=1}^{n_h} y_j^{(2)} w_{j1}^{(2)} = \sum_{j=1}^{n_h} \tanh \left(\sum_{i=1}^{n_{var}} x_i w_{ij}^{(1)} \right) w_{j1}^{(2)}$$

Где n_{var} и n_h количество нейронов в входном слое и в скрытом, соответственно, $w_{ij}^{(1)}$ вес между нейронном входного слоя i и нейронном скрытого слоя j , а $w_{j1}^{(2)}$ является весом синапса между скрытым слоем j и нейроном выходного слоя. В процессе обучения, сеть обучается на N тренировочных данных $\mathbf{x}_a = (x_1, \dots, x_{n_{var}})_a$, $a = 1, \dots, N$. Для каждого тренировочного события вычисляется и сравнивается выходной сигнал с желаемым выходным сигналом $\hat{y}_a \in [1, 0]$ (1 для сигнального события и 0 для фонового). Функция E , измеряющая разницу выходного сигнала с желаемым выходным сигналом называется функцией потерь. С помощью функции потерь можно оценить качество работы нейронной сети во время обучения. Чем меньше функция потерь, тем реже наша сеть ошибается. Именно к минимизации этой функции и сводится обучение. Функция потерь E определяется как

$$E(\mathbf{x}_1, \dots, \mathbf{x}_N | \mathbf{w}) = \sum_{a=1}^N E_a(\mathbf{x}_a | \mathbf{w}) = \sum_{a=1}^N \frac{1}{2} (y_{out,a} - \hat{y}_a)^2,$$

Где \mathbf{w} матрица весов сети. Процесс обучения — это движение по гиперповерхности функции потери, целью которого является поиск минимума. Для поиска минимума функции набор весов, который минимизирует функцию потерь можете быть найден с использованием метода градиентного спуска, при условии что функция потерь есть дифференцируемая функция относительно входных весов. Начиная с случайного набора весов $\mathbf{w}^{(\rho)}$, веса обновляются путем перемещения на небольшое расстояние в \mathbf{w} — пространстве в направлении $-\nabla_{\mathbf{w}} E$ где E уменьшается наиболее быстро

$$\mathbf{w}^{(\rho+1)} = \mathbf{w}^{(\rho)} - \eta \nabla_{\mathbf{w}} E,$$

Положительное число η является скоростью обучения. Веса связанные с выходным слоем обновляются следующим образом

$$\Delta w_{j1}^{(2)} = -\eta \sum_{a=1}^N \frac{\partial E_a}{\partial w_{j1}^{(2)}} = -\eta \sum_{a=1}^N (y_{out,a} - \hat{y}_a) y_{j1}^{(2)},$$

А веса связанные с скрытым слоем обновляются

$$\Delta w_{ij}^{(1)} = -\eta \sum_{a=1}^N \frac{\partial E_a}{\partial w_{ij}^{(1)}} = -\eta \sum_{a=1}^N (y_{out,a} - \hat{y}_a) y_{j,a}^{(2)} (1 - y_{j,a}^{(2)}) w_{j1}^{(2)} x_{i,a},$$

Использование градиентного спуска гарантирует нахождение только локального минимума функции. Однако, значение функции потерь в локальном минимуме является приемлемым решением, если выходная ошибка сети меньше заданной. При обучении использовалось 5 эпох с скоростью обучения $\eta = 0.02$.

Кросс-валидация

Для оценки качества построенной модели используют метод кросс-валидации на k блоках, которая разбивает имеющиеся данные на k блоков (рис. 4) .

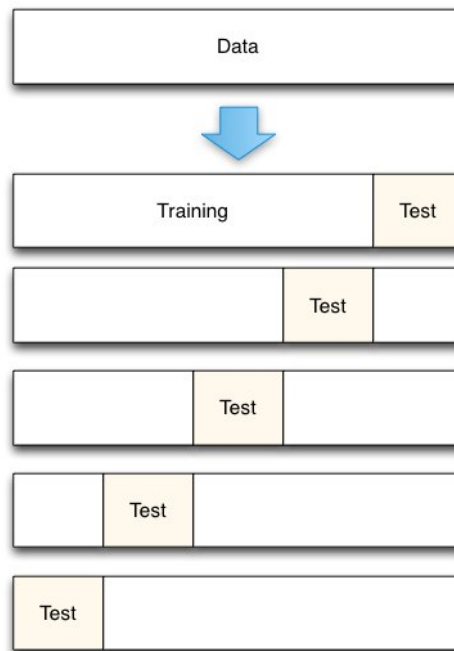


Рисунок 4 – Демонстрация метода кросс-валидация

Затем на $k - 1$ блоках данных производится обучение модели, а оставшаяся часть данных используется для тестирования. Данная процедура повторяется k раз, в результате чего, каждый из k блоков используется для тестирования. В результате получается оценка эффективности выбранной модели с наиболее равномерным использованием имеющихся данных.

3 Детектор ATLAS

ATLAS (от англ. A Toroidal LHC ApparatuS) — один из четырёх основных экспериментов на коллайдере LHC в Европейской Организации Ядерных исследований CERN в городе Женева (Швейцария) (рис. 5). Экс-

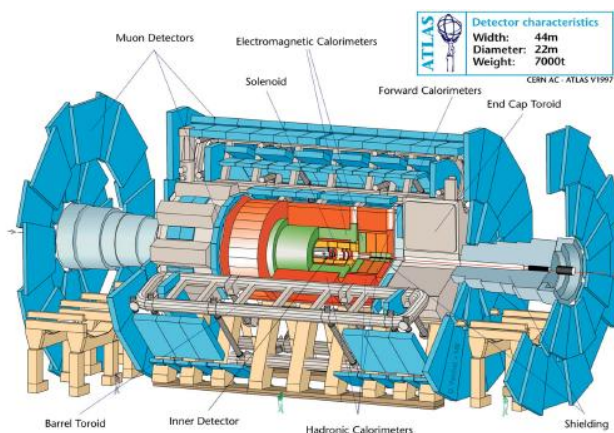


Рисунок 5 – Детектор ATLAS

перимент проводится на одноимённом детекторе, предназначенном для исследования протон-протонных столкновений. Детектор состоит из нескольких частей. Для восстановления треков и импульсов заряженных частиц используется внутренний детектор, окруженный сверхпроводящий магнитом, создающим магнитное поле 2 Тл. С помощью системы калориметров происходит измерение энергии частиц. На периферии детектора находится мюонный спектрометр, который необходим для измерения импульса мюонов. Для отбора событий используется система триггеров.

4 Используемые данные

Во время набора данных на эксперименте ATLAS рождается огромное число фоновых событий, для того чтобы понять какой вклад они вносят применяют статистический анализ, важным компонентом которого являются данные, сгенерированные методом Монте-Карло[5]. В работе использовались данные, полученные с использованием *xTauFramework*[1].

4.1 Экспериментальные данные

В работе использовались экспериментальные данные, набранные на детекторе ATLAS в 2015 году. Общая интегральная светимость 3.22 Фб^{-1} при столкновении протон-протонных пучков с суммарной энергией 13 ТэВ.

4.2 Монте–Карло моделирование

Смоделированные данные, используемые в работе, были получены методом Монте–Карло, с помощью генераторов Pythia[7] и Sherpa[4], и прошли всю цепочку реконструкций, на условии реальных протон-протонных столкновений эксперимента ATLAS сессия 2 (RUN 2). Каждому каналу соответствует свой уникальный номер DSID. Список Монте–Карло данных, использованных в работе, приведен в таблице 2.

Для сравнения Монте–Карло и реальных данных выполняется нормировка на светимость.

NOMINAL_pileup_combined_weight
lep_0_NOMINAL_MuEffSF_TTVA
lep_0_NOMINAL_MuEffSF_IsoFCTight
jet_NOMINAL_global_effSF_MV2c10
jet_NOMINAL_global_ineffSF_MV2c10
jet_NOMINAL_central_jets_global_effSF_JVT
jet_NOMINAL_central_jets_global_ineffSF_JVT
lep_0_NOMINAL_MuEffSF_Reco_QualMedium
lep_0_NOMINAL_MuEffSF_HLT_mu20_iloose_L1MU15_OR
HLT_mu50_QualMedium_IsoNone
jet_NOMINAL_forward_jets_global_effSF_JVT
jet_NOMINAL_forward_jets_global_ineffSF_JVT

Таблица 1 – Коррекционные коэффициенты

Для более точного согласия с распределениями из данных используются коррекционные коэффициенты, представленные в таблице 1, которые учитывают неточности в моделирование Монте–Карло и геометрию детектора. Коэффициенты коррекции, использованные в работе предоставляются Combined Performance (CP) Groups эксперимента ATLAS [2].

Sample	DSID	Generator	xs [pb]	k-factor
$W^+ \rightarrow \mu\nu$	361101	PowhegPythia8EvtGen	11500.9	1.01724
$W^- \rightarrow \mu\nu$	361104	PowhegPythia8EvtGen	8579.31	1.03576
$W^+ \rightarrow \tau\nu$	361102	PowhegPythia8EvtGen	11500.9	1.01724
$W^- \rightarrow \tau\nu$	361105	PowhegPythia8EvtGen	8579.31	1.03579
$Z \rightarrow \tau\tau$	361108	PowhegPythia8EvtGen	1950.63	1.02605
$Z \rightarrow \mu\mu$	361107	PowhegPythia8EvtGen	1950.63	1.02605
Top	410013	PhPy8EG_P2012	35.8455	1.054
Top	410014	PhPy8EG_P2012	35.8244	1.054
Top	410470	PhPy8EG_A14	452.346	1.13976
Top	410644	PowhegPythia8EvtGen	2.06146	1.017
Top	410645	PowhegPythia8EvtGen	128867	1.0167
Top	410646	PowhegPythia8EvtGen	35.8486	0.945
Diboson	363356	Sherpa_221_PDF30	2.20355	1
Diboson	363358	Sherpa_221_PDF30	3.4328	1
Diboson	363359	Sherpa_221_PDF30	24.708	1
Diboson	363360	Sherpa_221_PDF30	24.724	1
Diboson	363489	Sherpa_221_PDF30	11.42	1
Diboson	364250	Sherpa_221_PDF30	1.2523	1
Diboson	364253	Sherpa_221_PDF30	4.579	1
Diboson	364254	Sherpa_221_PDF30	12.501	1
Diboson	364255	Sherpa_221_PDF30	3.2344	1

Таблица 2 – Список каналов, использованных в генераторе Монте–Карло при моделировании данных.

4.3 Предварительный отбор

Для получения сигнального региона был проведен предварительный отбор. Кинематический отбор включает в себя ограничение на поперечный импульс P_T больше 27 ГэВ. Это значение было выбрано исходя из того, что планируется рассмотрение не только мюона, но и электрона, который имеет ограничение на поперечный импульс с учетом триггера в 27 ГэВ. Отбор на поперечную энергию E_T^{miss} больше 25 ГэВ, псевдобыстроту η по модулю меньше 2.47, поперечную массу M_T больше 40 ГэВ, определяемую как

$$m_T = \sqrt{2P_T(l)P_T(\nu)(1 - \cos(\phi(l) - \phi(\nu)))}, \quad (1)$$

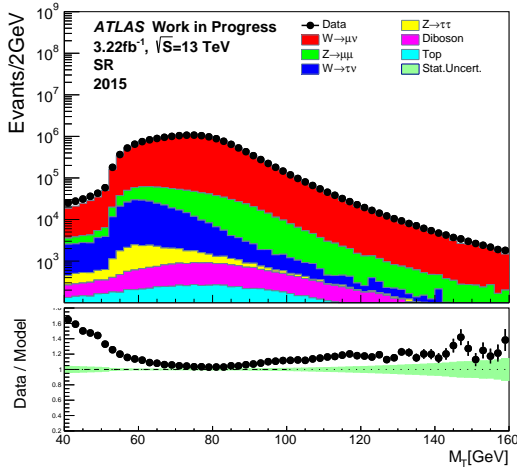
проводится для отсеечения большого количества КХД фона. Также использовалось ограничение на количество b -струй равное 0, количество лептонов в событии равное 1. Были применены критерии на идентификацию (Tight) и изоляцию мюона (Tight) (таблица 3).

Trigger	HLT_mu20_loose_L1MU15 muTrigMatch_0_HLT_mu20_loose_L1MU15 HLT_mu50 muTrigMatch_0_HLT_mu50
P_T	>27 ГэВ
E_T^{miss}	> 25 ГэВ
η	$ \eta < 2.47$
M_T	> 40 ГэВ
N лептонов	1
b струй	0
ID	Tight
Isolation	tight isolation
ptcone20/ P_T	< 0.6

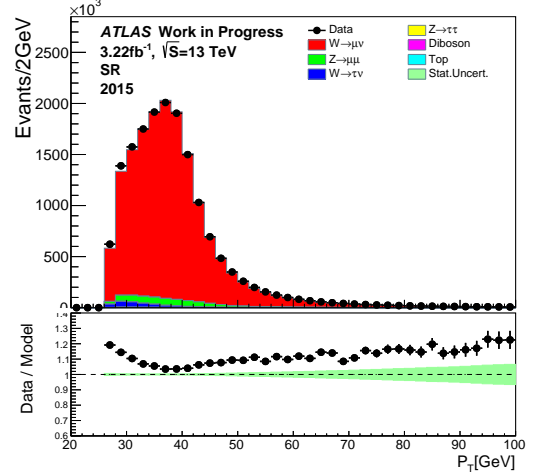
Таблица 3 – Предварительный отбор

4.4 Кинематические распределения

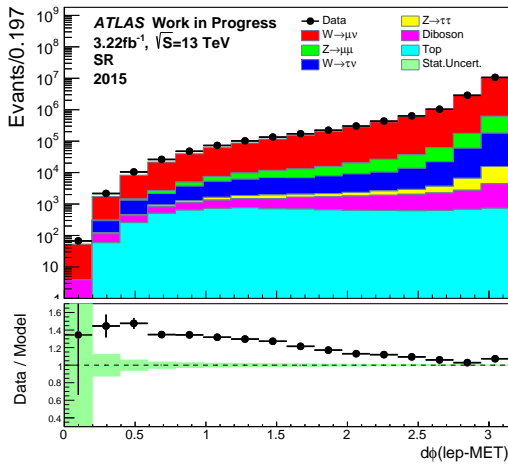
Было выполнено сравнение кинематических распределений для реальных и Монте-Карло данных (рис. 6 и 7), а также показано отношение реальных данных к данным Монте-Карло.



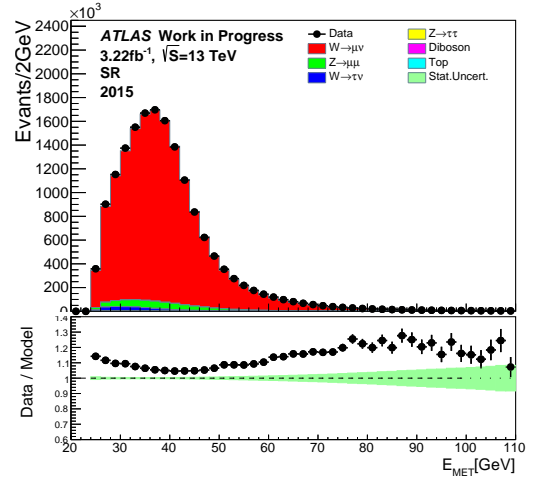
(а) Поперечная масса m_T



(б) Поперечный импульс P_T



(в) Разность углов ϕ потерянной энергии и лептона $d\phi(l - E_{\text{miss}})$

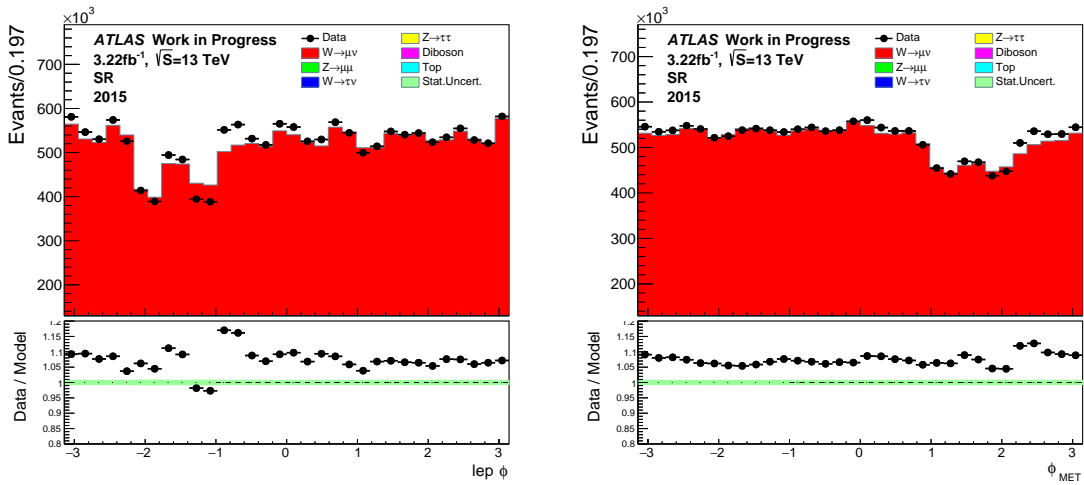


(г) Потерянная поперечная энергия E_{miss}

Рисунок 6 – Сравнение реальных и Монте-Карло данных для переменных: M_T , P_T , $d\phi(l - E_{\text{miss}})$, E_{miss}

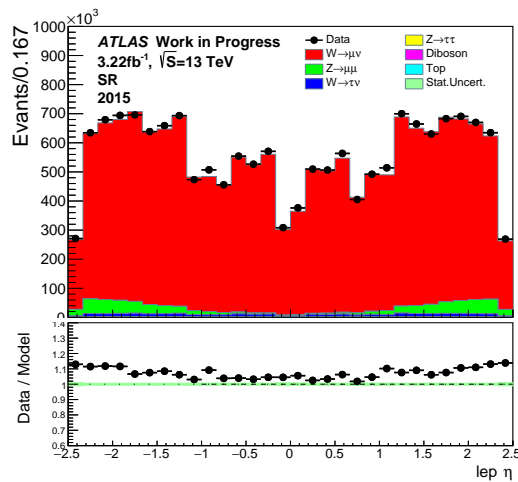
На представленном отношении отчетливо видно отклонение смоделированных Монте-Карло данных от реальных данных в мягкой части поперечного импульса P_T , потерянной энергии E_T^{miss} , поперечной массы M_T , а также при малой разности углов лептона и потерянной энергии $d\phi(\text{lep} - MET)$

. Одной из причин данных отклонений является неучтенный КХД фон, который не оценивался в данной работе. Отклонения в жесткой части по-



(а) Угол ϕ лептона

(б) Угол ϕ потерянной энергии



(в) Псевдобыстрота η

Рисунок 7 – Сравнение реальных и Монте–Карло данных для переменных: $\text{lep } \phi$, $\text{MET } \phi$, $\text{lep } \eta$

перечного импульса P_T , потерянной энергии E_T^{miss} , поперечной массы M_T связано с плохим моделированием Монте–Карло.

5 Обучение нейронной сети

В процессе обучения нейронной сети был проведен предварительный отбор. Для событий, использованных на этапе обучения, помимо отбора, указанного в главе 4.3, для повышения статистики сигнальных событий вместо отбора на идентификацию Tight был применен Medium. В каче-

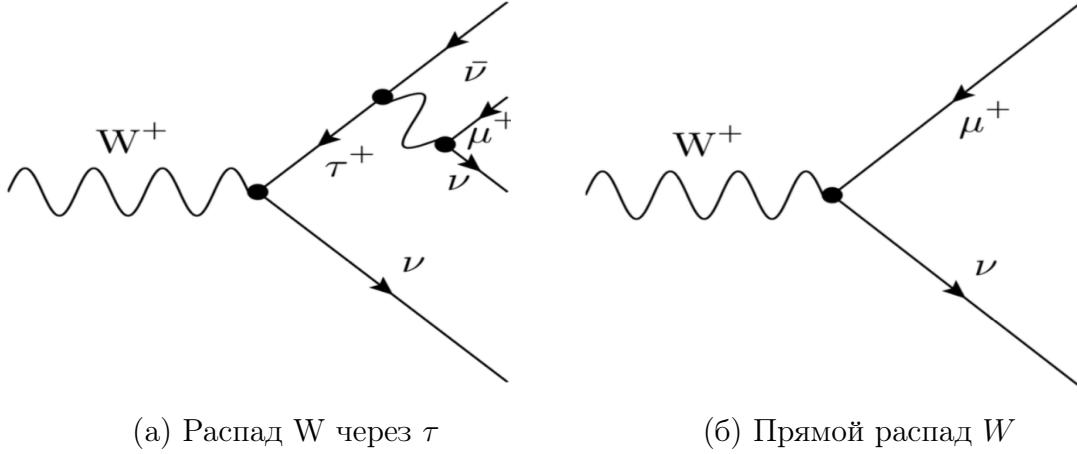


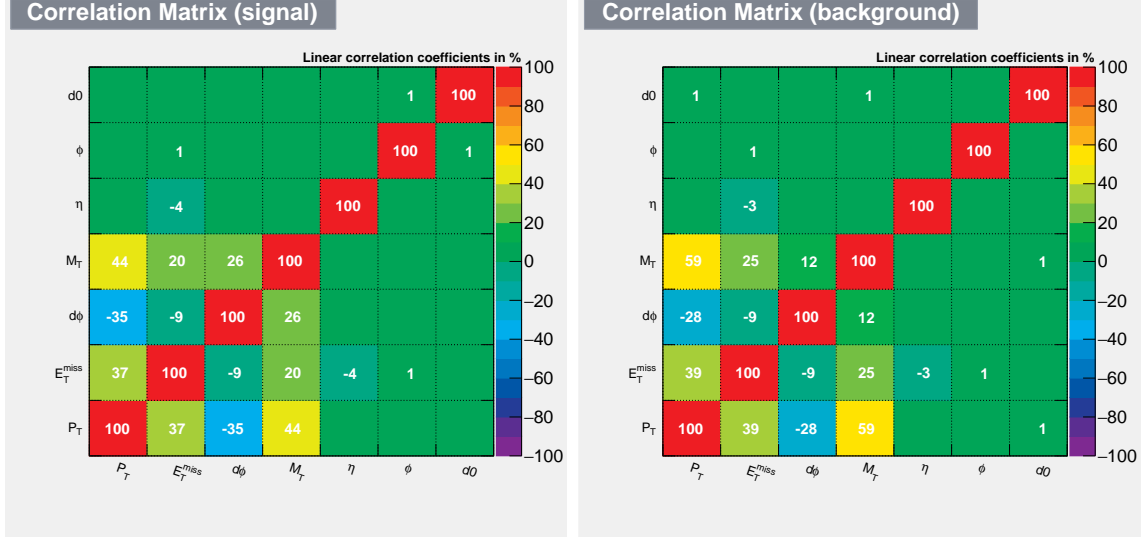
Рисунок 8 – Фейнмановские диаграммы

стве сигнала рассматриваются распады W в τ с последующим распадом в лептон второго поколения (рис. 8а), а в качестве фона рассматриваются прямые распады W в лептон второго поколения (рис. 8б). Полученная выборка для сигнального региона содержит 123148 событий, а в свою очередь полученная выборка для фоновых событий содержит 589390 событий. Соотношение тренировочных и тестовых данных было выбрано 70 на 30 процентов. В обучении нейронной сети использовались следующие переменные:

- Поперечная масса m_T .
- Поперечный импульс P_T
- Псевдобыстрота лептона η
- Угол потерянной энергии ϕ_{miss}
- Поперечная потерянная энергия E_{miss} .
- Разность углов ϕ для лептона и потерянной энергии $d\phi(lep - MET)$.

Тренировка нейронной сети проводилась без учета коррекционных коэффициентов, предоставленных Combined Performance (CP) Groups.

Были построены матрицы корреляции для всех переменных, используемых в обучении (рис.9).



(а) Матрица корреляции для сигнала (б) Матрица корреляции для фона

Рисунок 9 – Матрицы корреляции для сигнала и фона

Корреляция поперечной массы M_T с поперечным импульсом P_T для сигнальных событий составляет 44 процента, а для фоновых 59 процента. В свою очередь, корреляция поперечной массы M_T с потерянной энергии E_T^{miss} для сигнальных событий составляет 20 процентов, а для фоновых 25 процентов. Данный результат показывает ожидаемую зависимость поперечной массы от этих переменных (формула 1).

Прицельный параметр d_0 является дополнительным чувствительным показателем к идентификации непрямых распадов W в τ , так как τ лептон обладает временем жизни равным $2.9e-13$ с. Из этого следует, что при распаде W в τ мы имеем шире распределение для d_0 , чем при прямых распадах W в μ , поэтому эту переменную имеет смысл использовать как контрольную переменную в двумерном фитировании. Для проверки возможности использования прицельного параметра d_0 в двумерном фитировании выполнена проверка на корреляцию с остальными кинематическими переменными, судя по которой прицельный параметр d_0 не коррелирует ни с одной из переменных.

Чтобы оценить качество бинарной классификации, как правило используют кривую ошибок (ROC–integ)[8]. Чем больше интеграл под ROC кривой, тем лучше модель классифицирует события. В работе исследована зависимость значения кривой ошибки от используемых переменных при обучении (таблица 4).

Переменные	ROC–integ
$P_T, E_T^{miss}, d\phi(lep - MET), M_T$	0.732
$P_T, E_T^{miss}, d\phi(lep - MET), M_T, \phi_{MET}, \eta$	0.734
$P_T, E_T^{miss}, d\phi(lep - MET), M_T, \phi_{MET}, \eta, d0$	0.735

Таблица 4 – Зависимость кривой ошибок от используемых переменных

Для оценки построенной модели был использован метод кросс-валидации (таблица 5). Значение ROC–integ на каждом блоке не отличается от преды-

	Блок 1	Блок 2	Блок 3
ROC – integ	0.732	0.731	0.736

Таблица 5 – Результат кросс-валидации

дущего, и в целом нейронная сеть работает стабильно на каждом из блоков, что свидетельствует о независимости выборки данных при обучении нейронной сети.

Для определения вклада, который вносит каждая переменная при классификации, была посчитана значимость для каждой переменной, использованной при обучении нейронной сети (таблица 6). Наибольший вклад в

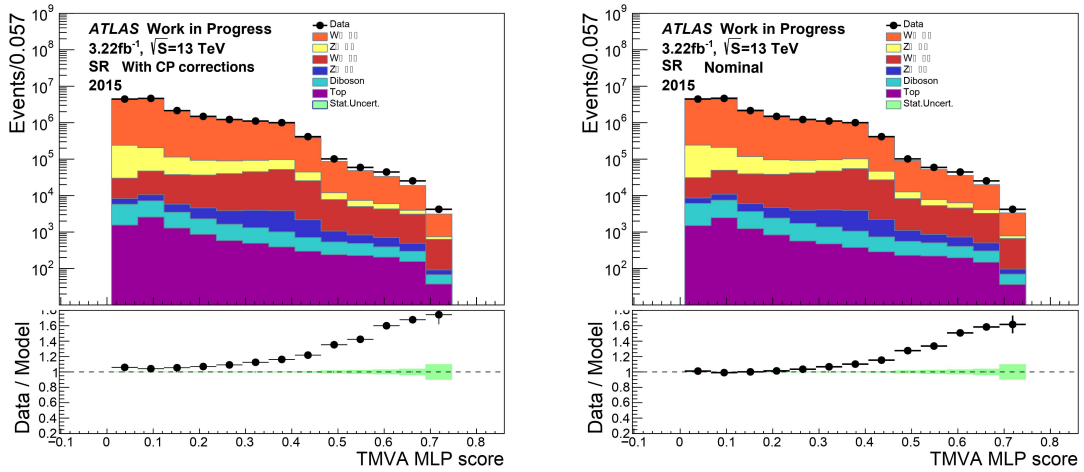
Переменная	P_T	M_T	$d\phi(lep - MET)$	E_T^{miss}	η	MET ϕ
Значимость	1.7e+03	4.7e+02	9.6e+00	2.7e+00	5.2e-02	2.5e-02

Таблица 6 – Значимость переменных

классификацию вносит поперечный импульс P_T и поперечная масса M_T , а наименьший псевдобыстрота лептона lep η и азимутальный угол для потерянной энергии MET ϕ .

6 Оценка отклика нейронной сети

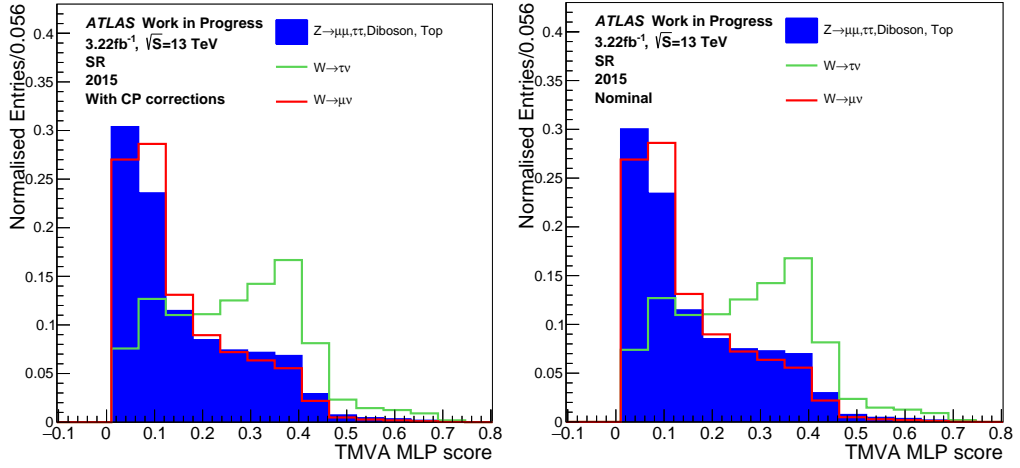
Готовая нейронная сеть использовалась для классификации данных указанных в главе 4. Классификация осуществлялась как с учетом коррекционных коэффициентов для данных Монте–Карло, так и без них (рис. 10а и 10б).



(а) Классификация с учетом коррекционных коэффициентов

(б) Классификация без учета коррекционных коэффициентов

Рисунок 10 – Отклик нейронной сети



(а) Классификация с учетом коррекционных коэффициентов (б) Классификация без учета коррекционных коэффициентов

Рисунок 11 – Сравнение отклика нейронной сети для трех основных физических процессов сигнального региона, после нормировки на интеграл

Были построены нормированные на интеграл распределения отклика для каждого канала Монте–Карло данных (рис 11а и 11б). На данных распределениях видно явное отличие формы распределения сигнала от фона, поэтому данная нейронная сеть может быть использована в двумерном фитировании. Для исследования влияния коррекционных коэффициентов на отклик нейронной сети, выполнено сравнение распределений отклика

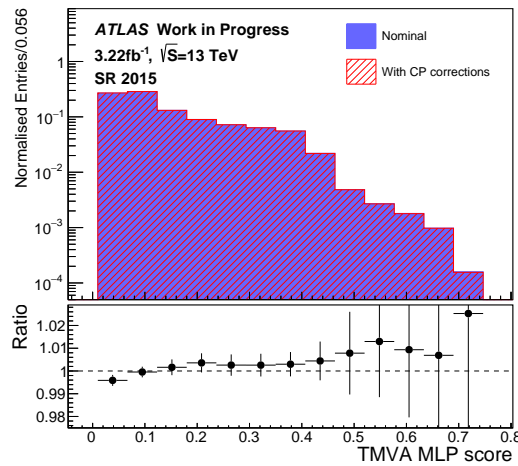


Рисунок 12 – Отношение классификаций

нейронной сети на Монте–Карло данных с учетом коррекционных коэффициентов и без, а так же построено их отношение (рис. 12).

По данному отношению видно, что отклик нейронной сети с использованием коррекционных коэффициентов отличается от отклика нейронной сети без коррекционных коэффициентов на 1-3 процента в области сигнального региона, что является критичным показателем для фитирования. Данный результат свидетельствует о том, что коррекционные коэффициенты меняют фазовое пространство сигнального региона, поэтому коррекционные коэффициенты стоит учитывать не только при классификации данных, но и на этапе обучения нейронной сети.

7 Заключение

В процессе выполнения данной научной работы были построены кинематические распределения для Монте–Карло данных и реальных данных 2015 года с энергией 13 ТэВ эксперимента ATLAS. Показано их расхождение, в том числе объясняемое неучтенным КХД фоном.

Была выполнена тренировка нейтронной сети, исследована зависимость ее точности работы от количества поданных на вход кинематических переменных. Получены значения интеграла под ROC кривой с использованием метода кросс валидации на 3 блоках, получена таблица значимости переменных, а также исследована корреляция переменных, использованных при обучении.

Выполнена оценка физических событий с помощью построенной нейронной сети, исследовано влияние коррекционных коэффициентов на ее отклик, показано сильное изменение в распределении отклика для Монте–Карло данных.

В процессе работы были получены начальные сведения о нейронных сетях и о объектно ориентированном программировании в языке Python.

В дальнейшем планируется тренировка нейронной сети с использованием корреляционных коэффициентов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. ATLAS xTauFramework software documentation. — URL: https://gitlab.cern.ch/Wlep_BR/xTauFramework/blob/master/doc/README.md.
2. Combined Performance (CP) Groups. — URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasPhysics>.
3. Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP / S. Schael [et al.] // Phys. Rept. — 2013. — Vol. 532. — P. 119–244.
4. Event Generation with Sherpa 2.2 / E. Bothmann [et al.] // SciPost Phys. — 2019. — Vol. 7, no. 3. — P. 034.
5. Monte Carlo event generators for high energy particle physics event simulation / S. Alioli [et al.]. — 2019. — Feb.
6. Precision measurement of the mass of the τ lepton / M. Ablikim [et al.] // Phys. Rev. — 2014. — Vol. D90, no. 1. — P. 012001.
7. *Sjostrand T., Mrenna S., Skands P. Z.* A Brief Introduction to PYTHIA 8.1 // Comput. Phys. Commun. — 2008. — Vol. 178. — P. 852–867.
8. TMVA - Toolkit for Multivariate Data Analysis / A. Hoecker [et al.]. — 2007.