

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

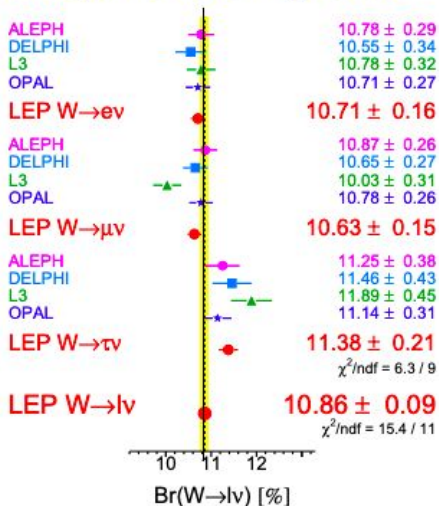
Научный руководитель: Пономаренко Д.Е.

Выполнил: Толкачёв Г.А.

Теоретическое введение

Лептонная универсальность

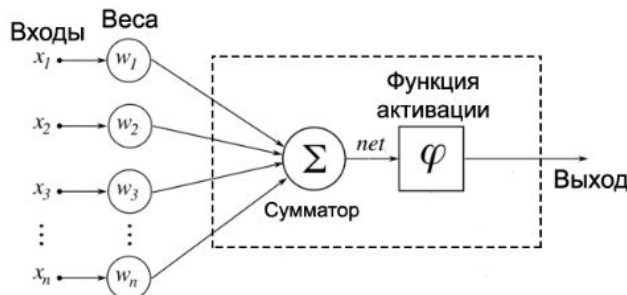
W Leptonic Branching Ratios



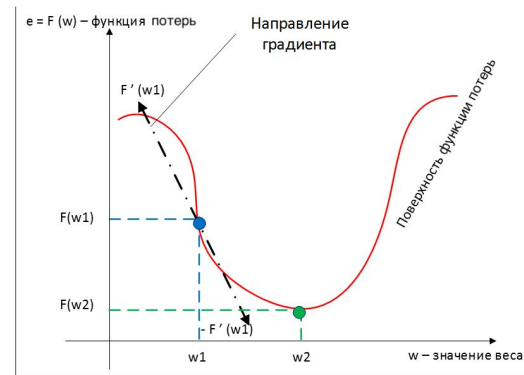
В результатах анализа с LEP имеются указания на отклонения от лептонной универсальности в отношениях сечения лептонного распада W в τ к полному сечению лептонного распада W бозона.

Нейронная сеть

Искусственная нейронная сеть - упрощенная математическая модель биологической нейронной сети. Она представляет из себя последовательность нейронов, соединенных между собой синапсами. Нейрон является базовым компонентом нейронной сети.



Метод градиентного спуска



Процесс обучения — это процесс подбора весов нейронной сети и движение по гиперповерхности функции потерь, целью которого является поиск минимума.

Данные и предварительный отбор

Экспериментальные данные были набраны на детекторе ATLAS в 2015 году. Общая интегральная светимость составляет 3.22 Фб^{-1} при столкновении протон-протонных пучков с суммарной энергией 13 ТэВ.

Монте–Карло данные получены с помощью генераторов Pythia и Sherpa.

Для сравнения Монте–Карло и реальных данных выполнялась нормировка на светимость. Для более точного согласия с распределениями из данных использовались коррекционные коэффициенты .

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

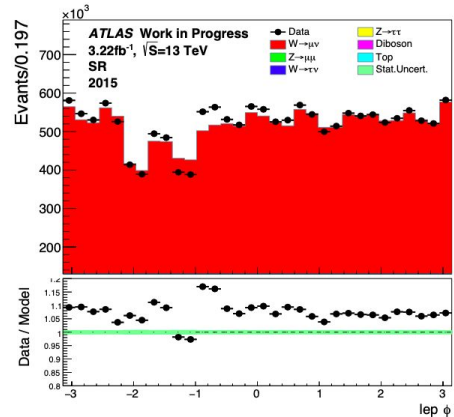
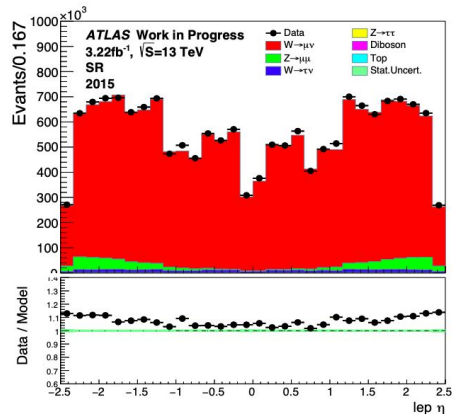
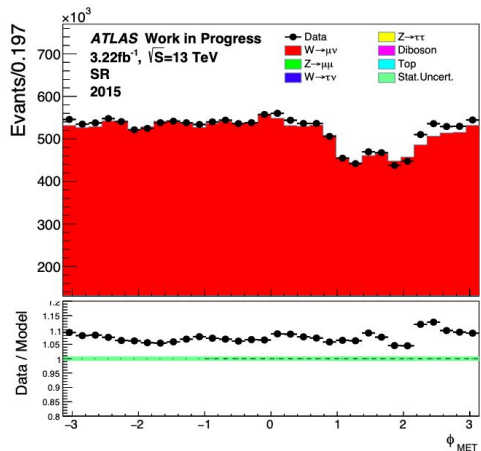
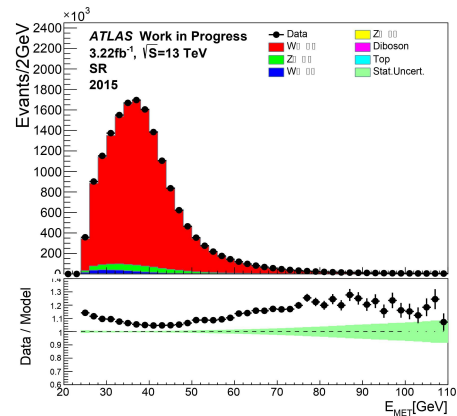
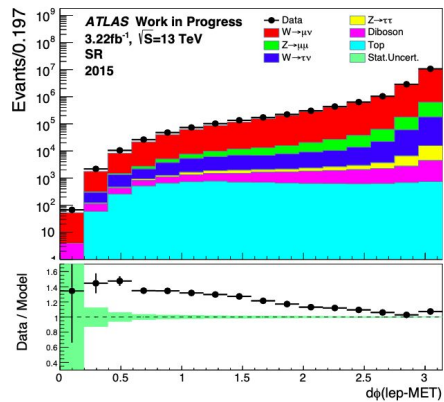
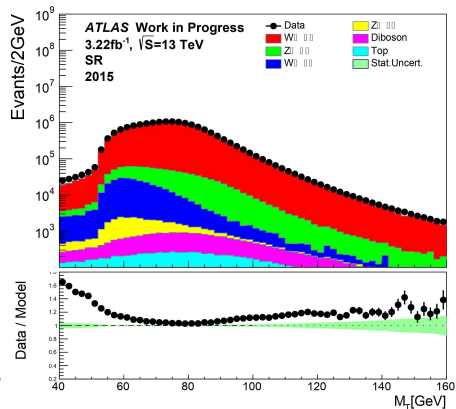
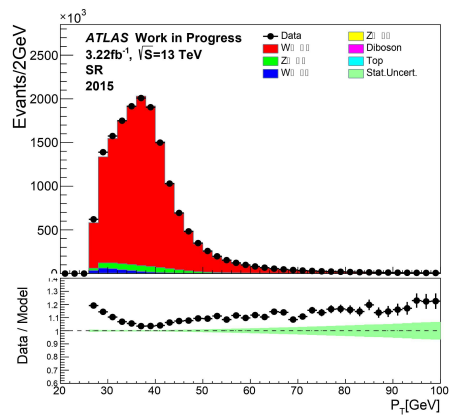
Sample	DSID	Generator	xs [pb]	k-factor
$W^+ \rightarrow \mu\nu$	361101	PowhegPythia8EvtGen	11500.9	1.01724
$W^- \rightarrow \mu\nu$	361104	PowhegPythia8EvtGen	8579.31	1.03576
$W^+ \rightarrow \tau\nu$	361102	PowhegPythia8EvtGen	11500.9	1.01724
$W^- \rightarrow \tau\nu$	361105	PowhegPythia8EvtGen	8579.31	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

```
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_loose_L1MU15_OR_HLT_mu50_QualMedium_IsoNone
jet_NOMINAL_forward_jets_global_effSF_JVT
jet_NOMINAL_forward_jets_global_ineffSF_JVT
```

Для получения сигнального региона был проведен отбор:

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

Кинематические распределения



Обучение нейронной сети

Кросс-валидация на 3 блоках

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

Значимость переменных использованных в обучении

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

Параметры нейронной сети:

Количество нейронов в

скрытом слое N+1

Количество эпох 5

Параметр шага градиентного

спуска 0.02

Тип функции активации

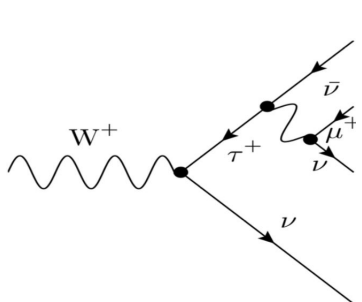
гиперболический тангенс

Соотношение тренировочных и

тестовых данных 70/30

Сигнальные события:

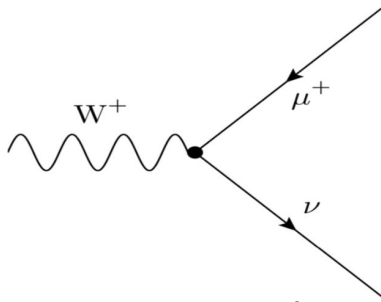
$W \rightarrow \tau + \nu \rightarrow \mu + \nu + \bar{\nu}$



Сигнальных событий
123148

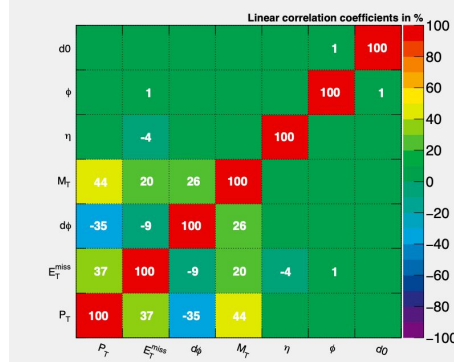
Фоновые события:

$W \rightarrow \mu + \nu$

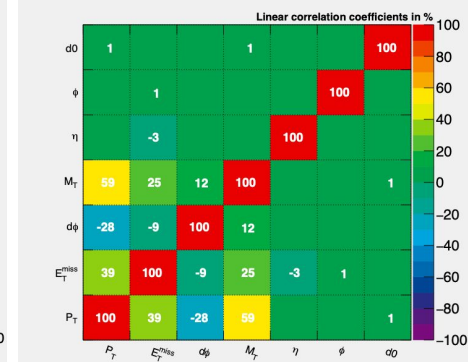


Фоновых событий
589390

Correlation Matrix (signal)



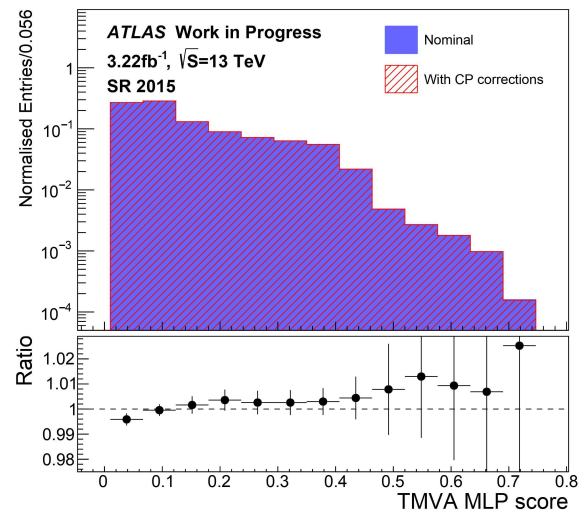
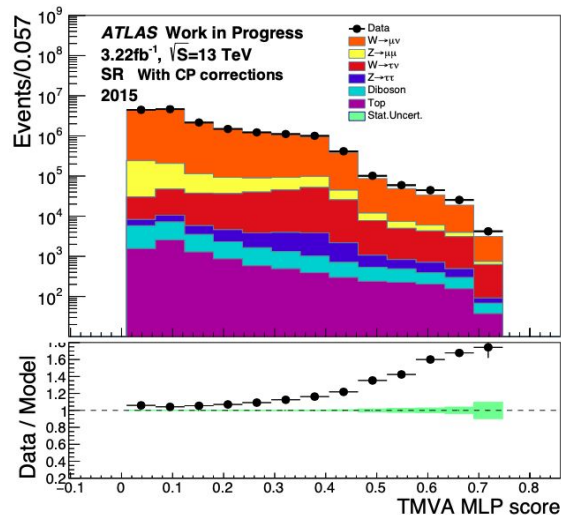
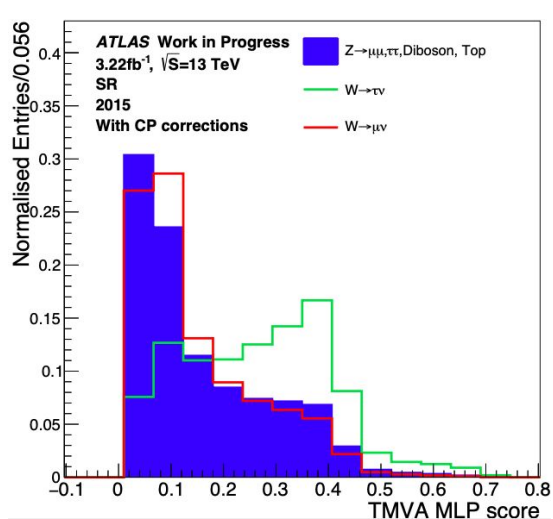
Correlation Matrix (background)



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

Прицельный параметр d_0 используется как контрольная переменная.

Оценка отклика нейронной сети



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 при обучении нейронной сети.

Изменение отклика нейронной сети на 1-3 процента в области сигнального региона.

Заключение

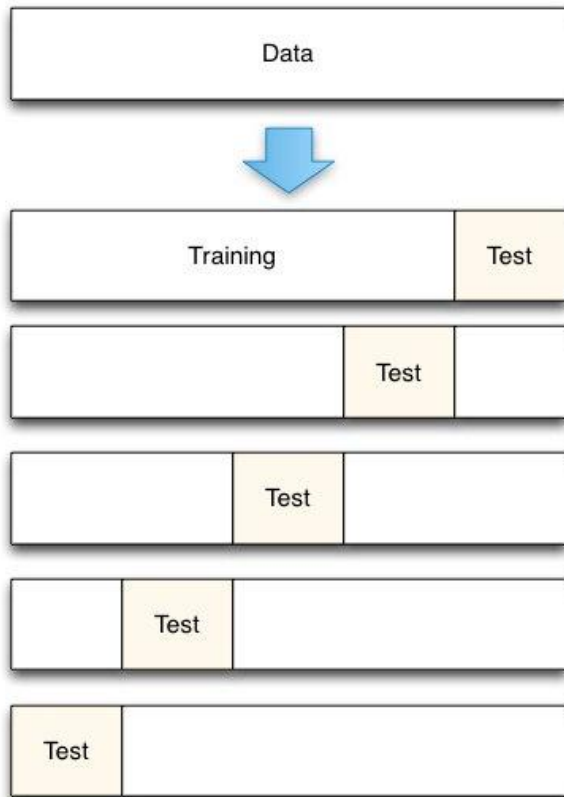
- Построены кинематические распределения для Монте-Карло данных и реальных данных 2015 года с энергией 13 ТэВ эксперимента ATLAS. Показано их расхождение, в том числе объясняемое неучтенным КХД фоном, а также из-за неточного моделирования P_t спектра лептонов и потерянной энергии.
- Выполнена тренировка нейронной сети, исследована зависимость ее 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 нейронной сети, исследовано влияние коррекционных коэффициентов на отклик нейронной сети, показано сильное изменение в распределении отклика для Монте-Карло.
- Произведено знакомство с объектно ориентированным программированием в языке Python.

В дальнейшем планируется:

- Тренировка нейронной сети с использованием корреляционных коэффициентов.

Спасибо за внимание!

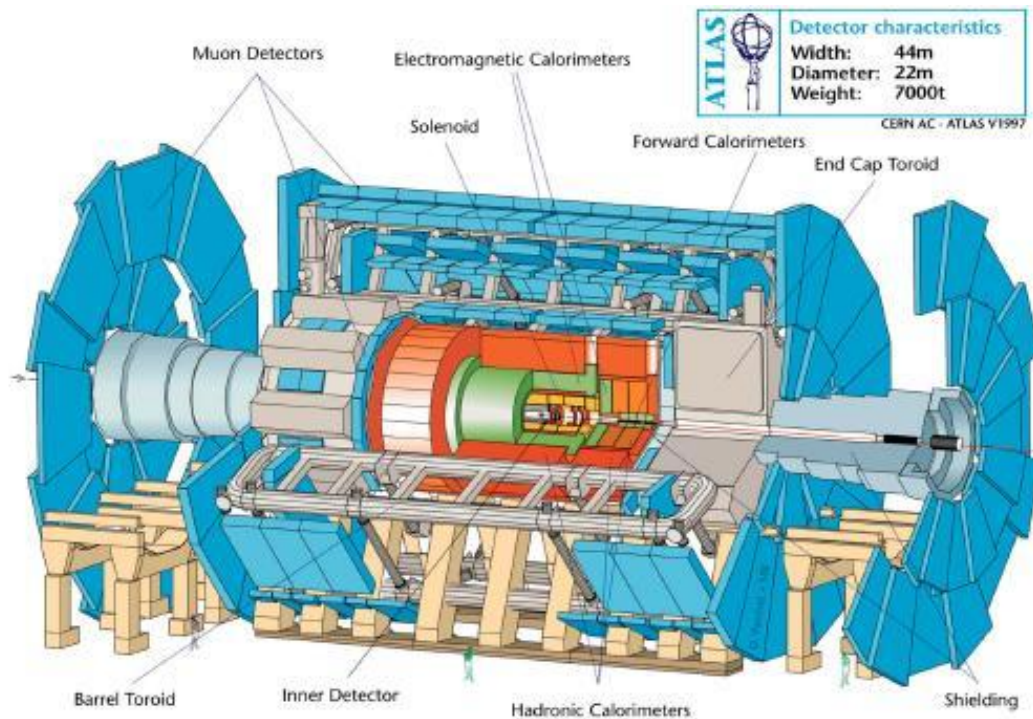
Кросс валидация



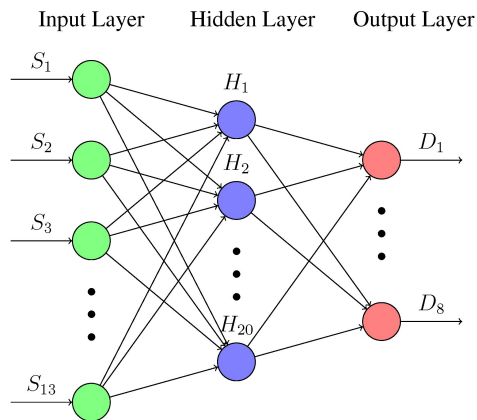
Для оценки качества построенной модели используют метод кросс-валидации на k блоках, которая разбивает имеющиеся данные на k блоков. Затем на $k - 1$ блоках данных производится обучение модели, а оставшаяся часть данных используется для тестирования. Данная процедура повторяется k раз, в результате чего, каждый из k блоков используется для тестирования. В результате получается оценка эффективности выбранной модели с наиболее равномерным использованием имеющихся данных.

Детектор ATLAS

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.



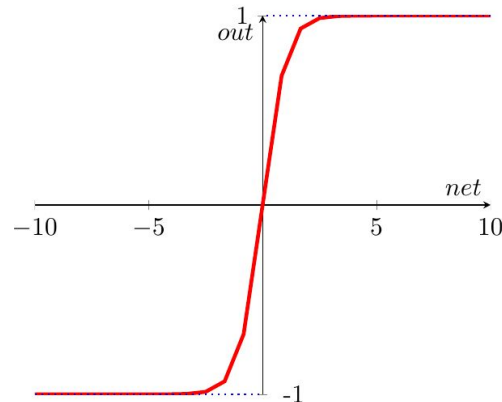
Нейронная сеть



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

Функция активации

Гиперболический тангенс



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
$P_T, E_T^{miss}, d\phi(lep - MET), M_T, \phi_{MET}, \eta, d_0$	0.735